

The reconstruction of 3-D marker trajectories in measuring heart motion

Citation for published version (APA):

Muijtjens, A. M. M. (1995). *The reconstruction of 3-D marker trajectories in measuring heart motion*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg.
<https://doi.org/10.26481/dis.19950210am>

Document status and date:

Published: 01/01/1995

DOI:

[10.26481/dis.19950210am](https://doi.org/10.26481/dis.19950210am)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

**The reconstruction of
3-D marker trajectories
in measuring heart motion**

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Muijtjens, Arnoldus Matheus Marie

The reconstruction of 3-D marker trajectories in measuring heart motion / Arnoldus Matheus Marie Muijtjens. - [S.l. : s.n.]

Proefschrift Maastricht. - Met lit. opg. - Met samenvatting in het Nederlands.

ISBN 90-9007955-6

Trefw.: binoculair zien / bewegingswaarneming.

Drukwerk: Datawyse / Universitaire Pers Maastricht

The reconstruction of 3-D marker trajectories in measuring heart motion

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Limburg te Maastricht,
op gezag van de Rector Magnificus Prof. dr. H. Philipsen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen op
vrijdag 10 februari 1995 om 16.00 uur

door

Arnoldus Matheus Marie Muijtjens

geboren te Maastricht in 1950

Promotores:

Prof. dr. ir. A. Hasman
Prof. dr. ir. T. Arts (Universiteit Twente)
Prof. dr. R.S. Reneman

Beoordelingscommissie:

Prof. dr. M.A. Allesie (voorzitter)
Prof. dr. ir. J.H. van Bommel (Erasmus Universiteit Rotterdam)
Prof. dr. A. Huson
Dr. ir. C.W.J. Oomens (Technische Universiteit Eindhoven)
Prof. dr. N. Westerhof (Vrije Universiteit Amsterdam)

Het verschijnen van dit proefschrift werd mede mogelijk gemaakt door de steun van de Nederlandse Hartstichting en de Stichting ter bevordering van de Research in de Cardiologie RESCAR. Bijdragen aan de promotiekosten werden eveneens verleend door Janssen Pharmaceutica BV, Medtronic Bakken Research Center BV, en Pie Medical Equipment BV.

Met dank aan collega's, familieleden en vrienden die, elk op hun eigen wijze, bijgedragen hebben aan de totstandkoming van dit proefschrift.

Aan mijn ouders
Aan Ineke, Nic en Cayke

Contents

Chapter 1	General Introduction	1
Chapter 2	Noise reduction in estimating cardiac deformation from marker tracks	15
Chapter 3	Extrapolation of incomplete marker tracks by lower rank approximation	23
Chapter 4	Tracking markers with missing data by iterative lower rank approximation	45
Chapter 5	3-D Reconstruction of marker positions from stereo images using ML estimation	65
Chapter 6	Establishing stereo correspondence for marker tracks	99
Chapter 7	General Discussion and Conclusions	127
	Summary	135
	Samenvatting	139
	Resumé	143
	Curriculum Vitae	145

General Introduction

Introduction

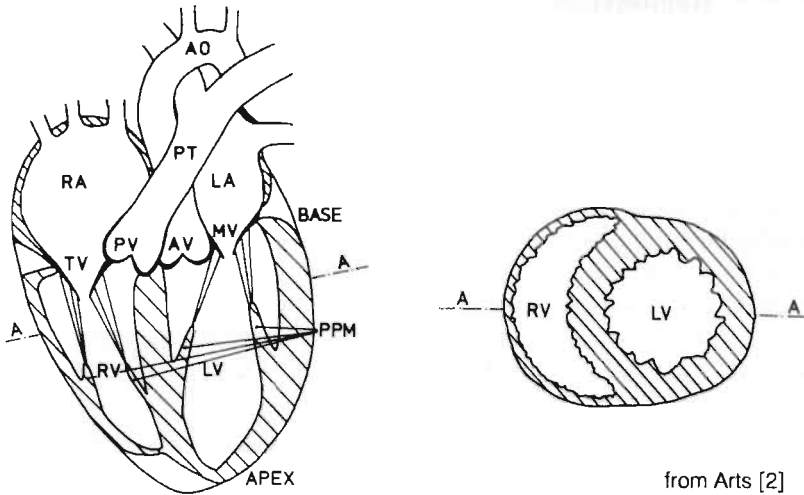
In this thesis methods are developed to reconstruct the 3-D trajectories of radiopaque markers which are attached to the heart. This general introduction starts with a brief description of the anatomy and function of the heart, followed by explaining the importance of measuring regional cardiac deformation. Methods currently used to measure regional deformation are briefly reviewed, and our choice for the technique of following radiopaque markers in an X-ray stereo camera set-up is motivated. The principle of measuring 3-D positions in an X-ray stereo camera set-up is presented, as well as the problems which occur when reconstructing the 3-D trajectories of many identical markers in a set-up of unknown geometry. In the developed solutions, a model for position estimation is used which is based on the Singular Value Decomposition [1] (SVD). The principle of the SVD model is briefly explained and the chapter concludes with a presentation of the aim and outline of the thesis.

Dynamic behaviour of the heart

The heart is a hollow, thick-walled muscle, which pumps the blood through the vascular system. The left panel of Fig. 1 schematically shows a long-axis cross-section of the heart. The heart consists of four cavities, the right (RA) and left (LA) atria and the right (RV) and left (LV) ventricles. The blood which returns from the systemic circulation is collected in the right atrium, and the blood from the pulmonary circulation is collected in the left atrium. The left atrium is separated from the left ventricle by the mitral valve (MV), and the right atrium from the right ventricle by the tricuspid valve (TV). During the systolic phase the blood in the left ventricle is ejected through the aortic valve (AV) into the aorta (AO). Similarly, the right ventricle empties via the pulmonary valve (PV) into the pulmonary trunk (PT).

All four aforementioned valves are located in the basal plane of the heart. The apex is located opposite to the base across the ventricles. In the right panel of Fig. 1 a schematic short-axis cross-sectional view of the heart is shown. The plane of cross-section is indicated in the left panel of Fig. 1 by the dot and dashed line AA. The free wall of the right ventricle is significantly thinner than that of the left ventricle. The inner and outer surfaces of the ventricles are the endocardium and epicardium, respectively. The tissue of the cardiac wall between the endocardium and the epicardium is referred to as the myocardium, and consists mainly of striated muscle.

The different phases of the cardiac cycle occur practically synchronously on the right and left side of the heart. So, for simplicity the phases are described with respect to the left half of the heart only. In the diastolic phase the left ventricle is filled with blood from the left atrium. Filling is completed by the atrial contraction, which is initiated by depolarization of the sinus node. Then the ventricle contracts and the systolic phase



from Arts [2]

FIGURE 1. *Left panel:* Long axis cross-section of the heart. Right atrium (RA), left atrium (LA), right ventricle (RV), left ventricle (LV), tricuspid valve (TV), mitral valve (MV), pulmonary valve (PV), aortic valve (AV), pulmonary trunk (PT), aorta (AO), papillary muscles (PPM). *Right panel:* Short axis cross-section of the heart in the plane indicated by AA.

starts. The pressure increase in the left ventricle causes the closure of the mitral valve. In this phase the two left ventricular valves are closed, so left ventricular volume remains constant. During this isovolumic contraction phase, left ventricular pressure increases rapidly, and when it exceeds aortic pressure, the aortic valve opens and the ejection phase starts. As a result of the relaxation of the heart the outflow stops and the aortic valve closes. The ventricular volume remains constant while ventricular pressure decreases in this so-called isovolumic relaxation phase. As a result of complete relaxation the atrium is filled and diastole starts again. When left ventricular pressure drops below left atrial pressure, the mitral valve opens and the rapid filling of the left ventricle starts.

Blood and oxygen is supplied to the myocardium by the coronary arteries. The usual consequence of stenotic atherosclerosis is an inadequacy of this supply caused by (partial) occlusion of these arteries. Regional myocardial ischemia is induced and myocardial function is restricted. Contraction of the ventricular wall is essential for the pump function of the heart. The reduced function in the ischemic region is reflected by reduced segment shortening or even segment lengthening [3-5]. In order to investigate cardiac mechanics in relation to ischemia, one must be able to accurately measure regional motion and deformation of the myocardium with sufficient resolution in space and time.

Regional measurement of cardiac deformation

Various techniques are in use to measure regional myocardial motion and deformation. These are based on two-dimensional echocardiography (2DE), magnetic resonance imaging (MRI) and following attached markers, visualized either optically or by the use of X-rays.

With ultrasonic 2DE the motion of the heart can be measured noninvasively [6-9]. The resolution of the method is limited, and distinct identifiable points cannot be followed in three dimensions. Recent developments in MRI are of importance. By a new method called tagging, specific regions of the myocardium are labelled and myocardial motion can be derived from the displacement and deformation of the patterns [10-14]. At the present state of the art, when using tagging in MRI, only the motion in the image plane direction is measured and the MR imaging procedure requires stable periodic motion during hundreds of heart cycles. Nevertheless, the rapid developments in this field are very interesting and may lead to an MRI-based non-invasive method of following distinct tagged points in 3-D with sufficient resolution in space and time.

Regional motion and deformation may be measured by following markers attached to the heart in a sequence of images. In our laboratory the deformation of the heart was investigated in open-chest dogs by following about 50 optical markers which were attached to the surface of the left ventricle [15-18]. Major drawbacks of the method are that an opened chest is required during measurements, only short term effects can be investigated (maximum duration experiment: one day), and the two-dimensional deformation of only a 40 mm \times 40 mm part of the surface of the anterior free wall of the L.V is measured.

In the coronary arterial tree, bifurcations can be visualized with cineangiography. They are used as natural markers when investigating 3-D heart motion in closed-chest experiments [3,19-21]. Radiopaque markers often have been implanted to serve as identifiable points in measuring heart motion, using single or stereo X-ray recordings of the heart [22-28]. At present, following the motion of implanted radiopaque markers by an X-ray stereo camera set-up seems to be most appropriate for obtaining accurate 3-D motion data of the heart with sufficient resolution in time and space [11,23] in closed-chest dog experiments. Moreover the X-ray stereo camera set-up offers good opportunities for automating detection and tracking of video recorded markers. Furthermore, the method is appropriate for detecting deformation over much longer periods than one cardiac cycle, for instance for detecting gradual cardiac dilatation or myocardial growth.

3-D position measurement in an X-ray stereo camera set-up

The principle of X-ray projection of radiopaque markers is schematically shown in the upper left panel of Fig. 2. Rays originating from the source point S_L pass through the

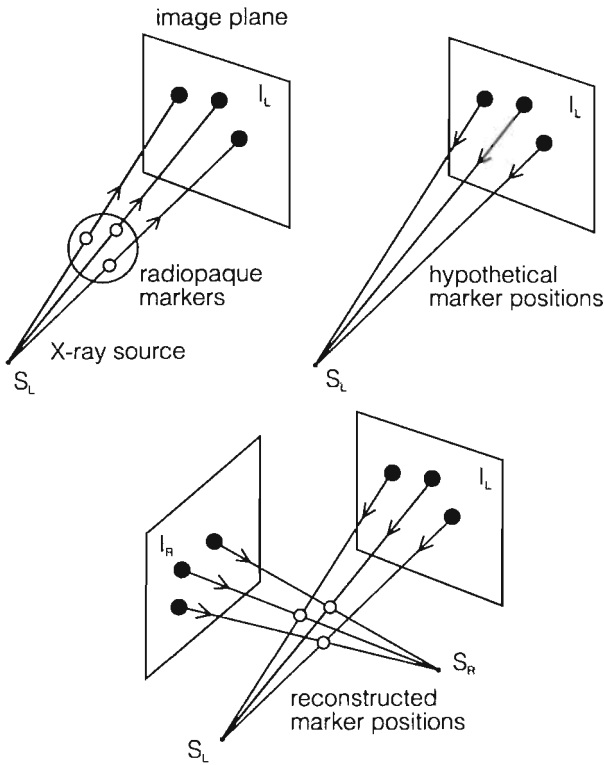


FIGURE 2. 3-D position measurement in an X-ray stereo camera set-up. *Upper left panel:* Shadow images (filled circles) are obtained of radiopaque markers (hollow circles) according to the central projection from source point S_L to the image plane I_L . *Upper right panel:* Marker images obtained with one camera contain information of only the direction of marker positions. *Lower panel:* When also images are obtained with a second camera (source point S_R and image plane I_R), the marker positions can be reconstructed at the intersection of corresponding imaging rays.

observed object and are detected in the image plane I_L . The intensity of the rays which pass the radiopaque markers (circles) is attenuated, and as a result corresponding X-ray shadow images (filled circles) are obtained in the image plane [29,30]. From an X-ray image recorded with one camera only the direction of the corresponding marker can be obtained by backprojection (upper right panel). When also images are obtained with a second camera, defined by source point S_R and image plane I_R , the location of the markers can be reconstructed at the intersections of the corresponding imaging rays (Fig. 2, lower panel).

Theoretically the marker images in both views lie in a so-called epipolar plane [31] defined by the two source points and the position of the corresponding marker. The restriction for stereo pairs to be located in an epipolar plane is called the epipolar constraint. In practice, due to noise the epipolar constraint holds only approximately. From the lower panel it also becomes clear that in case several markers are located in the same epipolar plane ambiguity results with respect to the correspondence of marker images in the two views.

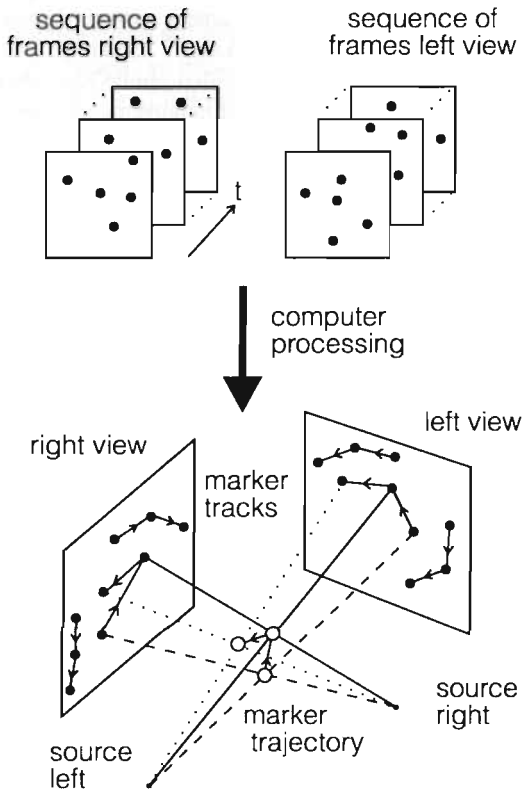


FIGURE 3. Reconstruction of 3-D marker trajectories. *Upper panel:* Sequences of frames (rectangles) obtained in two views, with detected marker image positions indicated (dots). *Lower panel:* Stereo camera set-up with marker tracks indicated in the image planes (connected dots), stereo pairs of imaging rays (solid, dotted and dashed lines), and reconstructed marker trajectory (connected open circles).

Reconstruction of 3-D marker trajectories

When a set of implanted radiopaque markers is observed in time by two cameras simultaneously in an X-ray stereo set-up, two sequences of video frames with marker images are obtained. After detection of the marker images, in each frame a set of marker image positions is obtained. The resulting data are presented in the upper panel of Fig. 3, with frames indicated by rectangles and marker image positions by dots. The marker trajectories can be reconstructed from these data as is shown in the lower panel. For each view the marker image positions in 3 consecutive frames are presented (dots). Sequences of marker image positions corresponding to the same marker, the so called

marker tracks, are indicated by the connecting lines and arrows. Three pairs of rays (dashed, solid, and dotted) are shown. Each pair corresponds to two simultaneously obtained images of the same marker obtained in the two views. When the geometry of the stereo camera set-up is known, the marker trajectory can be constructed by the sequence of 3-D marker positions at the intersections of the stereo pairs of rays.

The reconstruction procedure assumes the availability of undisturbed and complete marker image position data. In practice marker images may be missing because of detection failure, and when detected, in general their positions are disturbed by noise due to focal spot and motion blur [30], and due to image discretization [15].

In order to enable the reconstruction of complete, reliable and accurate 3-D marker trajectories, four types of information must be obtained:

- Position estimates for noise reduction and missing data substitution.
- Correspondence in time. Which marker image positions correspond to the same marker observed in consecutive frames.
- Stereo correspondence. Which marker image positions correspond to the same marker observed in different views.
- The geometry of the stereo camera set-up.

Model for position estimation using SVD

When the motion of a set of markers is observed by a camera in a sequence of frames to each marker corresponds a marker track, the sequence of marker image positions measured in time. The corresponding image position coordinates of each marker image as a function of time can be represented by a matrix which is schematically shown in the left panel of Fig. 4. When a large set of radiopaque markers (10-100) is attached to the heart wall, their movements are not independent. Then the corresponding set of marker image positions contains redundant information, which can be used to improve the position estimation. The motion of each marker image in such a case can be thought to be composed of a signal part and a noise part, with the signal part represented by a weighted sum of common motion components. When, as with SVD, the decomposition results in orthogonal motion components, the total energy in the matrix is the sum of the energy contributions per component. Then, generally, components with large energy contributions will represent common motion patterns associated with the signal part in the observations, while components with small energy contribution are associated with the independent noise patterns in the marker image tracks.

The estimation of the signal component of the marker tracks is based on the model shown in Fig. 4. Each observed track is written as the sum of a limited number of components with a proper amplitude coefficient 'c'. The shape of the first component is such that the maximum amount of energy of all observed signals is included. Of cour-

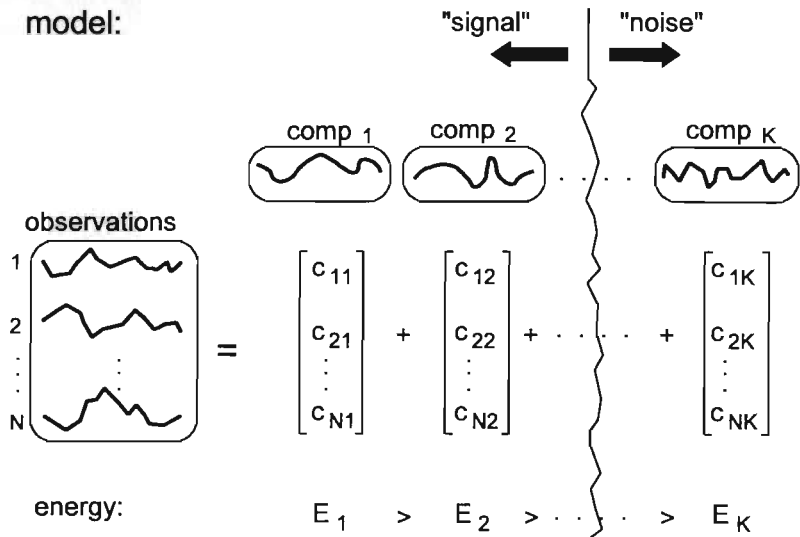


FIGURE 4. Schematic representation of the SVD model which is used to estimate the signal component in a set of marker tracks. The N observed time series are decomposed into K orthogonal components which are ordered according to decreasing energy contribution E to the set of observations. Generally, components with large energy contributions represent common motion patterns associated with the signal part, and components with small energy contributions represent the independent noise patterns in the tracks.

se, one component is not enough to describe all observations. After subtraction of the first signal component from the observations a residual is obtained. Similar to the first component, the shape of the second component is obtained as the one fitting best to the set of residual observations. This procedure is repeated, until the whole signal is decomposed. Inherently, the energy associated with each component decreases with increasing ordering number. Above a certain limit, the contribution of a component is below the noise level. At this point, the process is stopped. Then the complete set of observations can be described by a limited number of orthogonal components with their appropriate amplitude, and reduction of the noise or estimates of missing observations can be obtained.

Aim and outline of the study

Recently developed systems for measuring 3-D heart motion by following radiopaque markers [11,19,24,28] show some drawbacks which we try to avoid in our approach. The major drawbacks are:

- the use of a limited number of markers (14-22),
- incidental use of redundancy in the position data,
- the geometry of the stereo camera set-up is obtained in a separate calibration procedure by using a calibration object with a number of markers at known positions,
- stereo correspondences are obtained with the intervention of an operator in semi-automatic procedures.

The use of large sets of radiopaque markers enables more accurate measurement of differences in regional deformation in 3-D, but also results in more frequent occurrence of complications: (partly) overlapping marker images, crossing marker tracks, and ambiguities with respect to the identity of marker images in different views. However, with a more integral use of the redundancy in the image position data these complications can be overcome. More specifically, we hypothesize that

- the SVD model can be used to reduce the noise in the position data when a set of complete tracks is available, and a similar model can be used to estimate missing data or to extrapolate a set of partly known marker tracks (*Chapters 2 and 3*),
- by iteratively using the accurate extrapolations of partly known marker tracks which can be obtained by the latter application of the SVD model, the markers can reliably and automatically be tracked despite the complications caused by crossing tracks and missing marker image positions (*Chapter 4*),

In order to avoid a separate calibration procedure with a calibration object each time the set-up has been (accidentally) changed, calibration methods have been proposed which are based on using a set of known stereo correspondences of position data obtained during the measurements [32,33]. These methods describe the geometry of the stereo camera set-up by a linear model with a set of 8 parameters which is inherently redundant since a minimum of 5 parameters suffices to describe the geometry in a nonlinear model. Also, these methods assume that a set of stereo correspondences is available to be used as constraints in the calibration. But when - as is the case - large numbers of identical markers are used, however, obtaining a reliable set of stereo pairs in an uncalibrated set-up is a nontrivial task that has not been investigated. Nevertheless, we prefer a similar calibration procedure but with modifications for which we hypothesize that

- the calibration and 3-D reconstruction accuracy can be improved by using the parsimonious 5 parameter nonlinear model and by estimating the parameters according to

the maximum likelihood criterion. Besides, such a model is more appropriate to obtain insight into the estimation problem and into the structure of the reconstruction error (*Chapter 5*).

- a reliable set of stereo correspondences can be obtained iteratively in the uncalibrated set-up by basing stereo correspondence on matching tracks instead of single points. Furthermore, the noise influence and the computational effort can be reduced by calculating the mismatch for an approximate version of the 5 parameter model and by using only the most significant components of the SVD of the set of available marker tracks (*Chapter 6*).

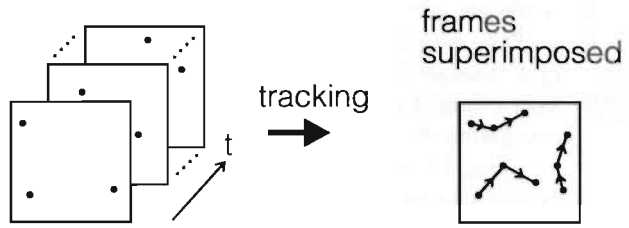
In computer simulations each developed method was extensively tested and evaluated, and performance limits were obtained. The SVD filtering method and the ILR extrapolation method (Chpts. 2 and 3) were evaluated with data obtained in the one camera optical set-up described in the Section ‘Regional measurement of cardiac deformation’. Simulated data were obtained by using a relatively simple model of cardiac deformation. In the Chapters 4 to 6 the major part of the evaluation was based on computer generated 3-D data of markers attached to a simulated moving left ventricle (LV). The motion of the LV was generated according to a kinetic model which uses 13 time varying parameters to describe several modes of motion and deformation in succession [34]. Realistic motion patterns were generated by basing the simulations on estimates of the kinetic parameters obtained from biplane cine recordings of 14 radiopaque myocardial markers in an animal experiment.

Fig. 5 schematically presents the processing steps used in our approach to automatically reconstruct the 3-D trajectories of a set of markers. At the left the obtained information and the corresponding chapter in the thesis are indicated. Note that the ordering of the chapters corresponds to the logical order of developing the methods, which differs from the order in the processing as shown in Fig. 5.

FIGURE 5. Processing steps in the automatic reconstruction of 3-D marker trajectories. At the left the obtained information and the corresponding chapters of the thesis are indicated. Processing starts with the construction of tracks from the unidentified sets of marker images in the sequence of frames per view. Then the noise can be reduced, missing data can be estimated, and stereo pairs can be obtained by selecting a set of matching tracks. The set of stereo pairs is used as constraints in the calibration of the stereo camera geometry, which completes the information needed to reconstruct the marker trajectories.

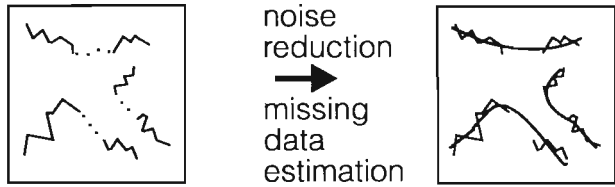
correspondence
in time

Chpt. 4



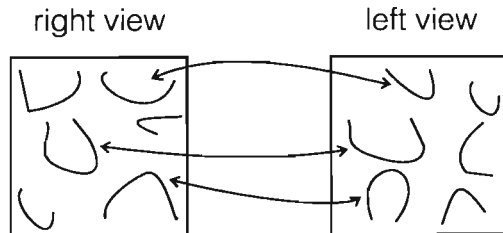
position
estimate

Chpt. 2,3



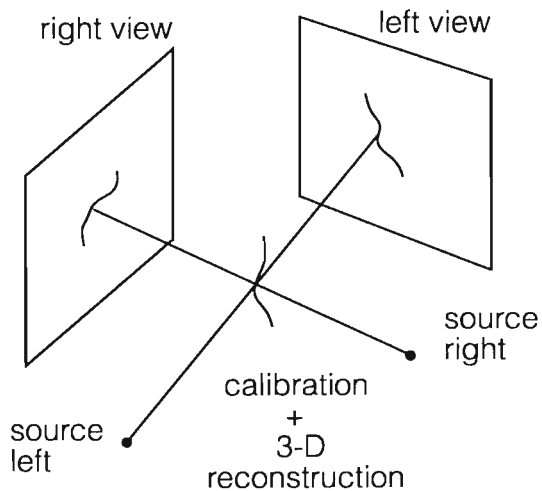
stereo
correspondence

Chpt. 6



stereo camera
set-up geometry

Chpt. 5



References

- [1] G.W. Stewart *Introduction to Matrix Computations*, (Academic, London, 1973).
- [2] M.G.J. Arts, *A mathematical model of the dynamics of the left ventricle and the coronary circulation*, Ph.D thesis, University of Limburg, Maastricht, The Netherlands, 1978.
- [3] Y. Kong, J.J. Morris and H.D. McIntosh, "Assessment of regional myocardial performance from biplane coronary angiograms," *Am. J. Cardiol.* **27**, 529-537 (1971).
- [4] F.W. Prinzen, T. Arts, A.P.G. Hoeks, and R.S. Reneman, "Discrepancies between myocardial blood flow and fiber shortening in the ischemic border zone as assessed with video mapping of epicardial deformation," *Europ. J. Physiol. (Pflügers Arch.)* **415**, 220-229 (1989).
- [5] H. Tomoike, Y. Urabe, S. Kaseda, and M. Nakamura, "Regional wall motion in the ischemic heart," In: M. Hori, H. Suga, J. Baan, and E.L. Yellin (eds.), *Cardiac mechanics and function in the normal and diseased heart*, pp. 257-262 (Springer, London, 1989).
- [6] M.P. Feneley and J.B. Hickie, "Validity of echocardiographic determination of left ventricular systolic wall thickening," *Circulation*, **70**, 226-232 (1984).
- [7] R.A. Levine, L.D. Gillam, and A.E. Weyman, "Echocardiography in cardiac research," In: H.A. Fozzard (ed.), *The heart and cardiovascular system*, pp. 369-452 (Raven, New York, 1986).
- [8] G.E. Mailloux, A. Bleau, M. Bertrand and R. Petitclerc, "Computer analysis of heart motion from two-dimensional echocardiograms," *IEEE Trans. Biomed. Eng.* **34**, 356-364 (1987).
- [9] T. Uchiyama, Y. Fujibayashi, Y. Sato, T. Sakamaki, and N. Kajiwarra, "Clinical application of echocardiographic imaging to diagnosis of coronary artery disease," *Japanese Circulation Journal*, **54**, 309-315 (1990).
- [10] E.A. Zerhouni, D.M. Parish, W.J. Rogers, A. Yang and E.P. Shapiro, "Human Heart: Tagging with MR Imaging - A method for noninvasive assessment of myocardial motion," *Radiology*, **169**, 59-63 (1988).
- [11] W.C. Hunter and E.A. Zerhouni, "Imaging distinct points in left ventricular myocardium to study regional wall deformation," In: J.H. Anderson (ed.), *Innovation in Diagnostic Radiology*, pp. 169-190 (Springer, London, 1989).
- [12] L. Axel and L. Dougherty, "MR Imaging of Motion with Spatial Modulation of Magnetization," *Radiology*, **171**, 841-845 (1989).
- [13] M.B. Buchalter, J.L. Weiss, W.J. Rogers, E.A. Zerhouni, M.L. Weisfeldt, R. Beyar and E.P. Shapiro, "Noninvasive quantification of left ventricular rotational deformation in normal humans using magnetic resonance imaging myocardial tagging," *Circulation*, **81**, 1236-1244 (1990).
- [14] P.B. Boesiger, S.E. Fischer, O.M. Hess, H.P. Krayenbühl, S.E. Maier and G. McKinnon, "Myocardial tagging for noninvasive quantification of myocardial motion," *MedicaMundi*, **36**, 106-110 (1991).
- [15] T.T. Prinzen, T. Arts, F.W. Prinzen and R.S. Reneman, "Mapping of epicardial deformation using a video processing technique," *J. Biomech.* **19**, 263-273 (1986).
- [16] F.W. Prinzen, C.H. Augustijn, T. Arts, M.A. Alessie and R.S. Reneman, "Redistribution of myocardial fiber strain and blood flow by asynchronous activation," *Am. J. Physiol.* **259**, H300-H308 (1990).

- [17] C.H. Augustijn, T. Arts, F.W. Prinzen and R.S. Reneman, "Mapping the sequence of contraction of the canine left ventricle," *Europ. J. Physiol. (Pflügers Arch.)* **419**, 529-533 (1991).
- [18] F.W. Prinzen, C.H. Augustijn, M.A. Allesie, T. Arts, T. Delhaas and R.S. Reneman, "The time sequence of electrical and mechanical activation during spontaneous beating and ectopic stimulation," *Eur. Heart J.* **13**, 535-543 (1992).
- [19] S.A. MacKay, M.L. Potel and J.M. Rubin, "Graphics methods for tracking three-dimensional heart wall motion," *Comput. Biomed. Res.* **15**, 455-473 (1982).
- [20] M.J. Potel, J.M. Rubin, S.A. MacKay, A.M. Aisen, J. Al-Sadir and R.E. Sayre, "Methods for evaluating cardiac wall motion in three dimensions using bifurcation points of the coronary arterial tree," *Invest. Radiol.* **18**, 47-57 (1983).
- [21] A.A. Young, P.J. Hunter and B.H. Smaill, "Estimation of epicardial strain using the motions of coronary bifurcations in biplane cinéangiography," *IEEE Trans. Biomed. Eng.* **39**, 526-531 (1992).
- [22] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers," *Circulation*, **52**, 859-867 (1975).
- [23] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard," *Circulation*, **61**, 966-972 (1980).
- [24] J.B. Garrison, W.L. Ebert, R.F. Jenkins, S.M. Yionoulis, H. Malcom, G.A. Heyler, A.A. Shoukas, W.L. Maughan and K. Sagawa, "Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms," *Comput. Biomed. Res.* **15**, 76-96 (1982).
- [25] B. Hannaford and S.A. Glantz, "Adaptive linear predictor tracks implanted radiopaque markers," *IEEE Trans. Biomed. Eng.* **32**, 117-125 (1985).
- [26] N.B. Ingels Jr, G.T. Daughters II, E.B. Stinson, E.L. Alderman and D.C. Miller, "Three-dimensional left ventricular midwall dynamics in the transplanted human heart," *Circulation*, **81**, 1837-1848 (1990).
- [27] A.S. Douglas, E.K. Rodriguez, W. O'Dell and W.C. Hunter, "Unique strain history during ejection in canine left ventricle," *Am. J. Physiol.* **260**, H1596-H1611 (1991).
- [28] M. A. Niczyporuk and D.C. Miller, "Automatic tracking and digitization of multiple radiopaque myocardial markers," *Comput. Biomed. Res.* **24**, 129-142 (1991).
- [29] B. Hallert, *X-ray Photogrammetry*, (Elsevier, New York, 1970)
- [30] C.H. Marshall, "Resolution, contrast, and noise in X-ray image recording," In: J.M. Taveras, and J.T. Ferrucci (eds.), *Radiology, diagnosis-imaging-intervention*, (Lippincott, London, 1991)
- [31] S.T. Barnard, and M.A. Fischler, "Computational stereo," *Computing Surveys*, **14**, 553-572 (1982).
- [32] H.C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, **293**, 133-135 (1981).
- [33] C.E. Metz and L.E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory," *Med. Phys.* **16**, 45-51 (1989).
- [34] T. Arts, W.C. Hunter, A. Douglas, A.M.M. Muijtjens, and R.S. Reneman, "Description of the deformation of the left ventricle by a kinematic model," *J. Biomech.* **25**, 1119-1127 (1992).

Noise reduction in estimating cardiac deformation from marker tracks

Arno M.M. Muijtjens, Jef M.A. Roos, Tim T. Prinzen*, Arie Hasman,
Robert S. Reneman**, and Theo Arts*

Departments of Medical Informatics and Statistics, Biophysics, and Physiology**,
Cardiovascular Research Institute Maastricht
University of Limburg, The Netherlands*

Published in Am. J. Physiol. 258 (Heart Circ. Physiol. 27): H599-H605, 1990.

Noise reduction in estimating cardiac deformation from marker tracks

A. M. M. MULIJTJENS, J. M. A. ROOS, T. T. PRINZEN,
A. HASMAN, R. S. RENEMAN, AND T. ARTS

*Departments of Medical Informatics and Statistics, Biophysics, and Physiology,
University of Limburg, 6200 MD Maastricht, The Netherlands*

MULIJTJENS, A. M. M., J. M. A. ROOS, T. T. PRINZEN, A. HASMAN, R. S. RENEMAN, AND T. ARTS. *Noise reduction in estimating cardiac deformation from marker tracks*. *Am. J. Physiol.* 258 (Heart Circ. Physiol. 27): H599-H605, 1990.—Deformation of the cardiac wall is measured by using optical or radiopaque markers attached to the wall. When digitized images are used, the accuracy of the measurement of a marker position is limited by pixel resolution and the size of the marker. The spatial accuracy is improved by singular value decomposition (SVD) filtering. This filtering procedure is based on the assumption that displacements of markers are mutually related because they are embedded in a common continuum. In a computer simulation with 48 markers in 51 video frames, the accuracy of the measurement of a marker position improved from 0.14 to 0.045 (SD) pixel. In an open-chest animal experiment, with markers on the surface of the heart, the deformation patterns were extracted more clearly using SVD filtering, while mutually related high-frequency components were not suppressed. In a 50-frame sequence of 256×256 video images of a $45 \text{ mm} \times 35 \text{ mm}$ deforming surface with 50 markers of 8 pixels in diameter, the marker position resolution improves from 0.1 to 0.03 (SD) pixel (6 μm). Strain is determined with an accuracy of 0.002 over a distance of 30 pixels (6 mm).

strain; shear; filter; singular value decomposition; image processing

THE MOTION OF MARKERS is often used to quantify mechanical properties of the object to which they are attached. The displacement of each marker is usually available in sampled form with respect to time. Deformation is determined by subtraction of neighboring marker positions. Inherently, such differentiating operations suffer from noise. Several mathematical techniques have been proposed to reduce the effect of noise, such as digital filtering (6), Chebyshev polynomial (6) or spline fitting (11), or using Fourier analysis (4). Although the results of these methods are generally presented for one-dimensional sequences, application to multidimensional data is possible. These methods have in common that the data are smoothed in time without taking advantage of a priori knowledge about the mechanism that generates the data.

In this paper we present a method for reducing noise in marker track coordinates. As a consequence, noise is reduced in derived-deformation parameters like strain and shear. The method is applied to experimental data that are used to map the deformation of the outer surface of the left ventricle in an open-chest animal preparation (9). About 50 white markers (diam 1.5 mm) are attached to the anterior free wall of the left ventricle over an area of $\sim 45 \text{ mm} \times 35 \text{ mm}$. Markers are detected and tracked by computer processing of the digitized video images (9).

The proposed filtering procedure is based on the assumption that the pattern of marker displacements can be composed of a limited number of basic time functions, like harmonics in Fourier analysis. In the filtering technique that we use, the shape of the basic functions is free and is fit best to the marker displacements found experimentally. Marker displacements are mutually related because they share a common continuum. This is particularly true if the myocardium is normally perfused, resulting in a homogeneous pattern of epicardial deformation. In case of local ischemia there is a distinct difference between the pattern of displacements of markers in the healthy and the ischemic region (8). The width of the border zone between ischemic, dyskinetic myocardium and normally perfused, well-contracting tissue is reported to be in the range of 10–20 mm (3, 7). Therefore, any method used to filter marker coordinates, which averages the time functions over neighboring positions, may limit the spatial resolution to a level of experimental significance.

The sampled marker coordinates as a function of time are stored in a rectangular position coordinate-time matrix. With the use of the singular value decomposition (SVD) of this noisy matrix, components associated with common displacements of markers are separated from those associated with random noise. A priori knowledge of the noise variance can be used to determine a threshold for the sum of squares associated with the noise. By truncating the SVD of the raw data matrix, the noise-free marker coordinate matrix is estimated in a least-squares sense.

MATHEMATICAL PRINCIPLE OF FILTERING BY SINGULAR VALUE DECOMPOSITION

The detected noisy position of M markers in N subsequent frames is represented by $2M \times N$ matrix \mathbf{U}

$$\mathbf{U} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M1} & x_{M2} & \dots & x_{MN} \\ y_{11} & y_{12} & \dots & y_{1N} \\ y_{21} & y_{22} & \dots & y_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ y_{M1} & y_{M2} & \dots & y_{MN} \end{pmatrix} \quad (1)$$

with $(x, y)_i$ the position of marker i in frame j , $1 \leq i \leq M$ and $1 \leq j \leq N$.

Any arbitrary $2M \times N$ matrix \mathbf{U} of rank r can be decomposed into a weighted sum of r $2M \times N$ matrices of rank 1 (10). In our case the number of frames generally is smaller than the number of marker coordinates. The matrix \mathbf{U} will be of full rank because of uncorrelated noise in the data. Hence \mathbf{U} is generally of rank N and its singular value decomposition is defined by

$$\mathbf{U} = \sum_{k=1}^N \lambda_k^{1/2} \mathbf{a}_k \mathbf{b}_k^T \quad (2)$$

with λ_k the k th eigenvalue of $\mathbf{U}^T \mathbf{U}$ in descending order, \mathbf{a}_k the k th unit-length $2M$ -dimensional eigenvector of $\mathbf{U} \mathbf{U}^T$, and \mathbf{b}_k the k th unit-length N -dimensional eigenvector of $\mathbf{U}^T \mathbf{U}$ ($\lambda_k^{1/2}$ is called the k th singular value of \mathbf{U}).

The decomposition can be interpreted as follows. The i th row of \mathbf{U} , representing one marker coordinate as a function of time, is a linear combination of basic time functions defined by the row vectors \mathbf{b}_k^T ($k = 1 \dots N$). The related amplitude factors are represented by the column vector \mathbf{a}_k , the element a_{ik} being associated with the i th marker coordinate. The row vectors \mathbf{b}_k^T are comparable to the harmonic sine waves with Fourier transform. However, the function \mathbf{b}_k^T is an optimal adaptation to the problem, which is generally not a set of sine waves.

The product of \mathbf{a}_k and \mathbf{b}_k^T is a $2M \times N$ matrix of rank 1. The eigenvectors \mathbf{a}_k , for $k = 1, 2, \dots, N$, are mutually orthogonal as well as the vectors \mathbf{b}_k , $k = 1, 2, \dots, N$. Because of this orthogonality, for the sum of squares of the elements of \mathbf{U} it holds that

$$\sum_{i=1}^{2M} \sum_{j=1}^N U_{ij}^2 = \lambda_1 + \lambda_2 + \dots + \lambda_N, \quad \lambda_1 > \lambda_2 > \dots > \lambda_N \quad (3)$$

The squared singular value, λ_k , corresponds with the sum of squares (energy) attributed to the \mathbf{b}_k^T components in the rows of matrix \mathbf{U} .

The larger singular values are associated with common components in the displacements of markers sharing the same continuum. The smaller singular values are asso-

ciated with less energy and probably with marker displacements which are more random and less coherent. In the present filtering technique, the common components are considered to be the signal, while the random, incoherent components represent noise which should be eliminated. So matrix \mathbf{U} is filtered to matrix \mathbf{R} by skipping the components with the smaller energy contributions

$$\mathbf{R} = \sum_{k=1}^L \lambda_k^{1/2} \mathbf{a}_k \mathbf{b}_k^T \quad (4)$$

The matrix \mathbf{R} defined as above is the best rank L approximation of \mathbf{U} in least-squares sense. Appropriate truncation, i.e., a proper choice of L , is obtained by using an estimated level of uncorrelated noise as a threshold. Noise energy E due to error variance σ^2 in each measurement of a marker coordinate is

$$E = 2M \cdot N \cdot \sigma^2 \quad (5)$$

When the decreasing number of degrees of freedom of the residual sum of squares with increasing value of L is taken into account, an estimate of L is found by minimizing the squared difference D^2 with

$$D = \left(\sum_{k=L+1}^N \lambda_k \right) - [2M \cdot N - (2M + N - L) \cdot L] \cdot \sigma^2 \quad (6)$$

The first part of the right-hand side represents the observed residual sum of squares, being the cumulative eigenvalues of the skipped components, as a function of L . The second part, which is used as the threshold, is the expected residual sum of squares as a function of L assuming the residuals to be uncorrelated random noise with variance σ^2 .

Note that in an experimental setup Eq. 6 can also be used to estimate σ^2 if the number L of relevant signal components is known on the basis of other, independent information. In that case D should be set to zero.

IMPLEMENTATION

In our data set generally $N < 2M$, so the eigenvalues λ_k and eigenvectors \mathbf{b}_k of matrix $\mathbf{U}^T \mathbf{U}$ are calculated first. Then the vector \mathbf{a}_k can be calculated by using $\mathbf{U} \mathbf{b}_k = \lambda_k^{1/2} \mathbf{a}_k$. After substitution of the chosen value of σ^2 , the cutoff value L is determined according to Eq. 6 and the filtered version of \mathbf{U} is calculated according to Eq. 4. The filter program is written in FORTRAN on a VAX 11/780 computer. Subroutines from the Numerical Algorithms Group FORTRAN library (5) are used to calculate the matrix $\mathbf{U}^T \mathbf{U}$ (F01CLF) and the eigenvalues and eigenvectors of $\mathbf{U}^T \mathbf{U}$ (F01AJF followed by F02AMF).

PERFORMANCE

Three experimental conditions were used to study the performance of the SVD-filtering procedure: computer simulation of a deformed surface, measurement in a calibration setup, and measurement in an animal experiment.

Computer simulation. In a computer evaluation, the measurement of epicardial deformation of a partially

ischemic left ventricular wall is simulated. The 256×256 image contains 48 markers (Fig. 1). Deformation in the left (24 markers, $x \leq 112.5$) and the right (24 markers, $x > 112.5$) halves are considered to be associated with normal and ischemic areas, respectively. The position of the i th marker as a function of time is represented by

$$\begin{bmatrix} x_i(t) \\ y_i(t) \end{bmatrix} = \begin{bmatrix} x_i(0) \\ y_i(0) \end{bmatrix} + f(t) \begin{bmatrix} a_{xx} & a_{xy} \\ a_{xy} & a_{yy} \end{bmatrix} \begin{bmatrix} x_i(0) \\ y_i(0) \end{bmatrix} \quad (7)$$

with $a_{xx} = 1$, $a_{xy} = -0.5$, and $a_{yy} = 0.5$.

The function $f(t)$ equals $f_1(t)$ in the normal area and $f_2(t)$ in the ischemic area (Fig. 2, bottom, solid lines 1 and 4, respectively). The shapes of these functions are schematic representations of physiological measurements performed earlier in our laboratory. The simulation defined by Eq. 7 generates marker positions as a function of time containing only two basic patterns, $f_1(t)$ and $f_2(t)$. Hence, the simulated noise-free data matrix has exactly three nonzero singular values associated with the mean coordinate value of a marker and the two different deformation patterns, respectively. In the simulation, uncorrelated gaussian noise (variance 0.02 pixel²) is added to the marker coordinates.

To show how spatial resolution is preserved, we determined deformation on the basis of noise-free as well as filtered simulated data in four overlapping areas aligned perpendicularly to the border between the two deformation fields (Fig. 2). The discontinuity is found to occur for $>90\%$ between the centers of the areas 2 and 3 on both sides of the border, indicating that spatial smoothing due to filtering is minor.

In an analysis of noise suppression properties, deformation parameters are estimated from the noisy marker displacements in 16 regions each containing 3 markers

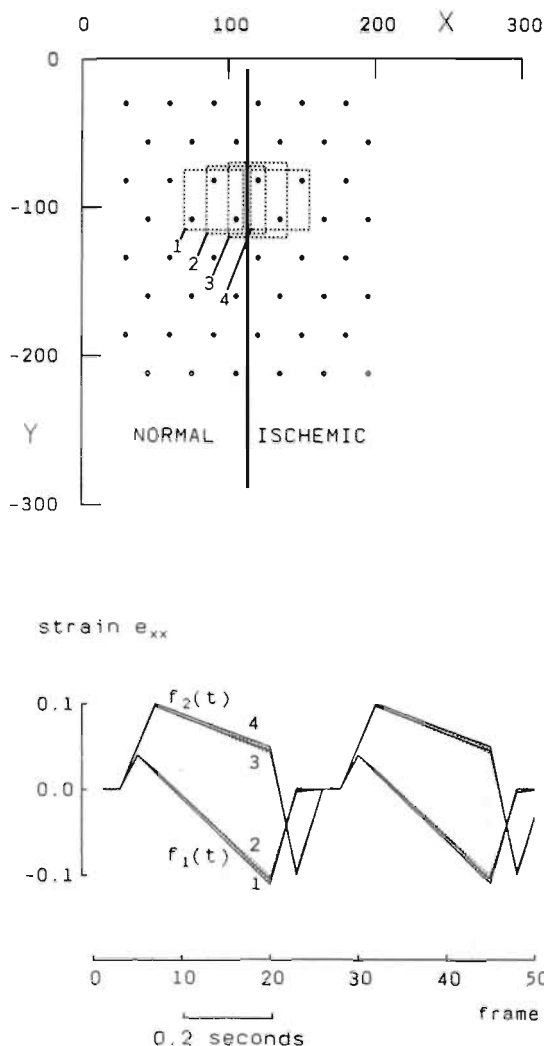


FIG. 2. Strain gradient of simulated data at border line. Top: 4 subregions at border. Bottom: strain e_{xx} as a function of time in subregions shown in top. Solid lines, noiseless strain; dotted lines, strain corresponding to SVD-filtered data. Solid lines 1 and 4 correspond to noiseless strain within the normal [$f_1(t)$] and ischemic [$f_2(t)$] areas.

(9). In Fig. 3A, for both areas four noisy tracings of e_{xx} thus calculated are shown (dotted lines). Figures 3B–D show the same deformation signals after SVD filtering of the marker coordinates with 51, 26, and 13 frames, respectively. The associated cumulative eigenvalues are shown in Fig. 4 (solid lines) together with the applied threshold levels, which are calculated according to Eq. 6 (dotted lines). Obviously, in this example the first three components represent the signal. The threshold level separates signal and noise correctly. Furthermore, unlike most filtering techniques, the frequency spectrum of the signals is quite unaffected as is shown by the preservation

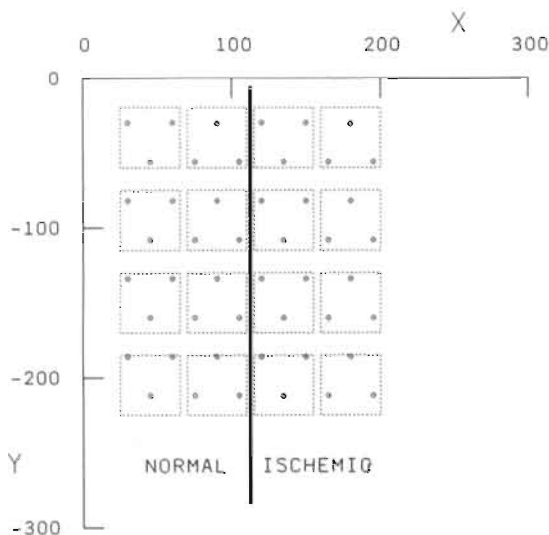


FIG. 1. Computer-generated grid of marker positions used in the simulations. Dotted squares, subregions in which deformation is calculated. Bold vertical line, border between simulated normal and ischemic areas.

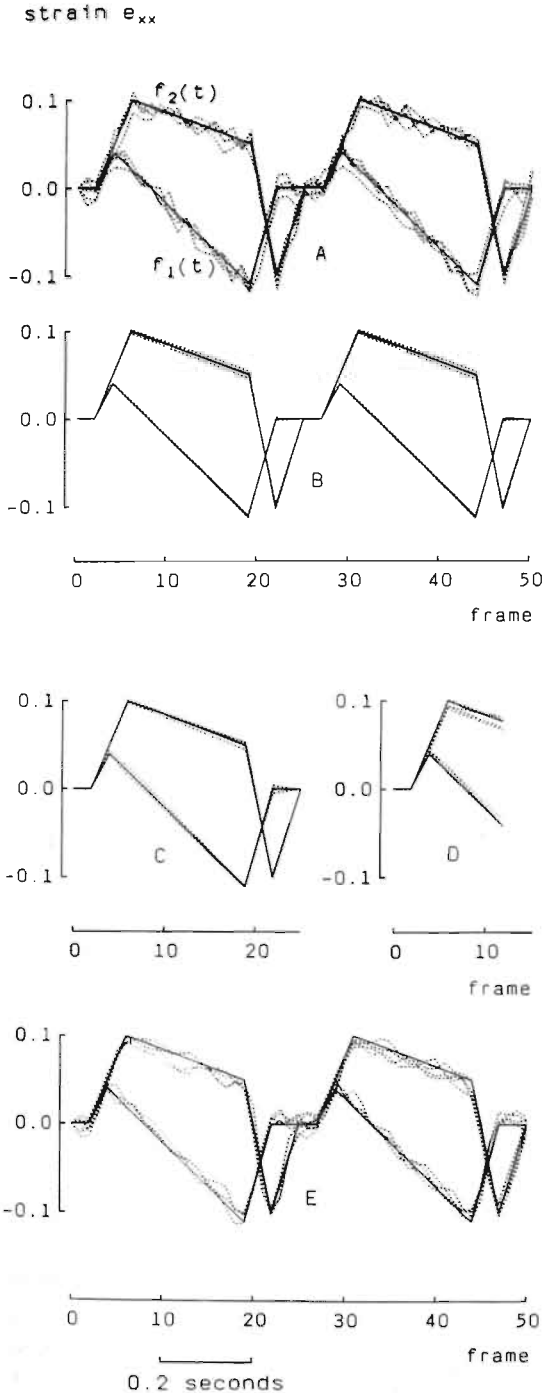


FIG. 3. Simulation of strain e_{xx} as a function of time. A: solid lines, noiseless strain in the normal [$f_1(t)$] and ischemic [$f_2(t)$] areas. Dotted lines, strains in upper 8 subregions of Fig. 1, as calculated on the basis

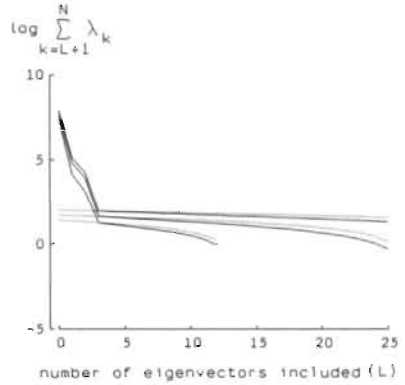


FIG. 4. Logarithm of cumulative eigenvalues (solid line) and filter threshold curve (dotted) according to Eq. 6 for the simulation shown in Fig. 3A. Curves are shown for 51 (top), 26 (middle), and 13 (bottom) frames filtered, respectively.

TABLE 1. Simulation: SD of noise in SVD-filtered data relative to unfiltered data (100%)

Number of Frames Filtered	Marker Coordinates (x,y)	Strain Parameters (e_{xx}, e_{yy}, e_{xy})
13	48%	45%
26	37%	24%
51	30%	18%

of sharp corners and straight parts of the signals $f_1(t)$ and $f_2(t)$. Table 1 illustrates the reduction of noise in the marker coordinates due to the filtering.

For comparison, the simulated data were also smoothed by low-pass filtering of the time sequences of the marker coordinates. We used the low-pass filter of module 1T of the statistical software package BMDP (2). The cutoff frequency of the filter (20.3 Hz) was chosen to yield a reduction of power similar to the reduction after SVD filtering of these data. As can be seen in Fig. 3, low-pass filtering reduces noise poorly (E) as compared to SVD filtering (B). When lowering the cut-off frequency, more power is lost and the time functions are significantly distorted by smoothing.

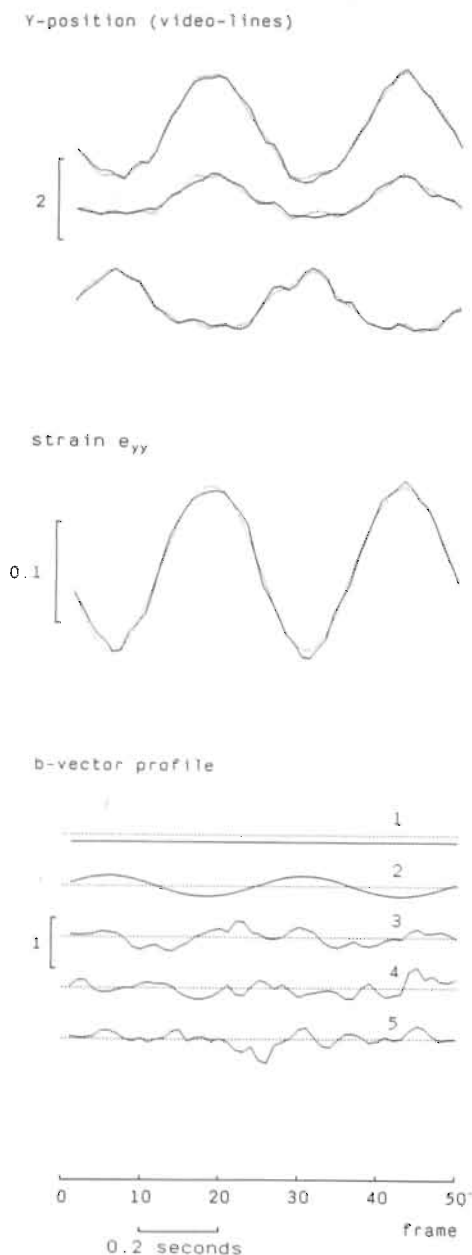
Measurement in a calibration setup. In a calibration setup 69 markers are distributed more or less randomly over a flat plate. The plate performs a 2-Hz, weakly damped oscillatory rotation around an axis parallel to the x direction aligned with the horizontal axis of the camera. The tilting angle of the plate oscillates around 45° with the optical axis of the camera and with an amplitude of $\sim 5^\circ$. Because of the use of a teleobjective (200-mm focal length), perspective distortion due to changes in distance to the camera is minimized. Thus strain $e_{yy}(t)$ varies sinusoidally with minor nonlinear distortion.

Analysis starts with preprocessing the raw data to

of unfiltered noisy marker position data (error variance 0.02). B: strain signals obtained by singular value decomposition filtering of noisy position data (51 frames filtered). C: similar to B for 26 frames. D: similar to B for 13 frames. E: effect of low-pass filtering; compare with B.

reduce the correlated noise patterns, which are caused by the interlacement of the lines in subsequent even and odd video frames. In this procedure the marker coordinates in subsequent frames are averaged. As a consequence, the error variance of the coordinate data is halved.

In this setup the SVD filter is evaluated by comparison



of noise levels in raw and filtered time courses of the y -coordinate data of the markers and the derived deformation parameter e_{yy} . The top panel of Fig. 5 shows the y coordinate as a function of time for three markers in the neighborhood of the rotational axis. The middle panel shows the corresponding e_{yy} . By passing only 3 of the 50 possible components, the filter yields (dotted lines) close approximations of the y coordinates resulting in a smooth, sinusoidal $e_{yy}(t)$. Extending the number of passed components provides only minor improvements of the fit to the experimental data. The use of only three components is also consistent with the patterns of the basic functions b_k^T , as shown in Fig. 5, *bottom*. The first three basic functions have a clearly visible systematic and periodic pattern, whereas the other ones are noisy. Thus setting L equal to 3 and D equal to 0 in Eq. 6, the error variance σ^2 is estimated to be 0.007. This value of σ^2 is used in the calculation of the value of L for the filtering of the data obtained in physiological experiments. After all, the camera setup in these calibration experiments is also used in the physiological experiments, so one may expect the errors in these measurements to be similar to those in the calibration data.

Measurement in an animal experiment. The SVD-filtering procedure has also been applied to data obtained in an animal experiment. The animal preparation, data acquisition, and data processing have been described in detail before (9). In addition, similar to the procedure used for the calibration data, the coordinates in subsequent frames have been averaged to reduce the correlated noise caused by the interlacement of video lines.

The experiment was performed in an open-chest dog preparation. About 50 white markers were attached to the epicardial surface and their motion was recorded on videotape. Markers were detected and tracked by computer processing of the digitized images. Marker position data from two heartbeats were filtered with the SVD method: one heartbeat in the control situation and one 2 min after induction of local ischemia by occluding the descending branch of the left interventricular coronary artery (Fig. 6). The coordinate data of 35 markers in 39 frames were filtered. The threshold was calculated by using Eq. 6 with $\sigma^2 = 0.007$ in accordance with the results of the calibration setup.

Strain along the epicardial fiber direction was estimated in five regions containing at least three markers. These regions were situated along a line crossing the interface between normally perfused (Fig. 6, *bottom*, lower tracings) and ischemic regions (*bottom*, upper tracings) during total coronary artery occlusion. Under control conditions, epicardial fibers shorten in all areas during the ejection phase. During coronary artery occlusion shortening is diminished considerably in the ischemic area. The filtered results (dotted lines) are smooth compared with the unfiltered signals (solid lines), but the frequency content is still high.

FIG. 5. Results obtained in calibration setup with an approximately sinusoidal variation of strain e_{yy} . *Top*: y coordinates of 3 markers are shown as a function of time. Solid lines, raw data; dotted lines, the filter output for $L = 3$. *Middle*: strain e_{yy} associated with the data in top. *Bottom*: functions b_k^T , which correspond to the 5 largest eigenvalues λ_k (dotted lines: base lines).

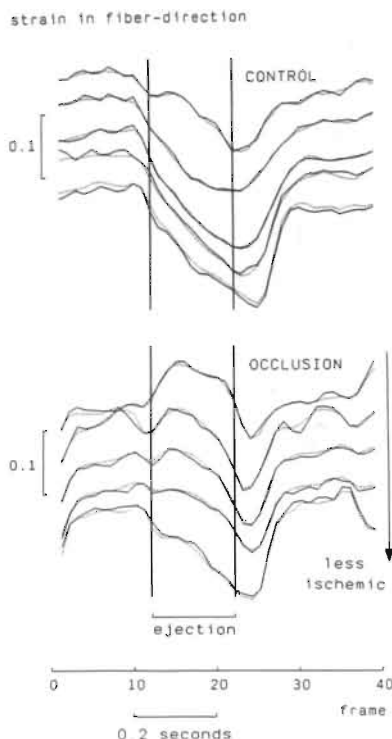


FIG. 6. Myocardial fiber strain measured at epicardium of free wall of left ventricle in control situation (top) and during occlusion of the anterior descending branch of left coronary artery (bottom), in 5 subregions from normal (lower tracing, bottom) to ischemic (upper tracings, bottom). Distance between subregions is ~ 7 mm. Solid lines correspond to raw data, dotted lines to filtered data. So that the different tracings may be distinguished, they have been shifted vertically. Vertical lines denote beginning and end of ejection phase based on recording of aortic volume flow.

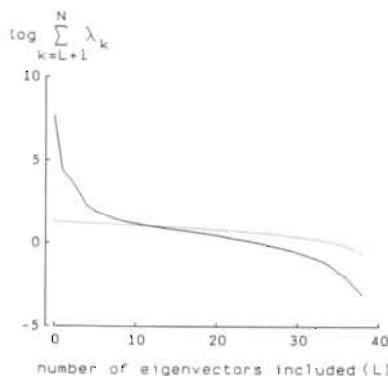


FIG. 7. For experimental data associated with Fig. 6, figure shows the logarithm of cumulative eigenvalues (solid line) and filter threshold values (dotted line) according to Eq. 6, with error variance chosen equal to 0.007. Only curves for occlusion data are shown; results for control data are similar.

Figure 7 shows the cumulative eigenvalues (solid line) and the filter threshold level (dotted line) as a function of L for the situation with occlusion. Compared with the simulations (Fig. 4), the transition from signal to noise is more gradual. This is not surprising, since the real data are expected to contain more than just three distinct signal patterns, as is the case for the simulated data. The threshold yields for L a value of 13 (control) and 12 (occlusion). Thus data reduction is considerable, since there is a total of 39 possible components. Despite the presence of larger inhomogeneities in the field of deformation in cases of coronary artery occlusion, the number of components needed appears not to be more than needed for the more homogeneous control situation (Fig. 6).

DISCUSSION

The present study shows that with the use of SVD filtering, marker coordinates in a digitized video recording can be determined accurately. The effective resolution of changes in marker position can be improved by a factor of 8 over the actual pixel grid size. Thus standard video systems are able to carry information at a resolution of $2,000 \times 2,000$ positions for 50 markers of 8 pixels in diameter recorded in 50 frames. Contrary to low-pass filtering or polynomial fitting, time and spatial frequency are not severely limited by SVD filtering. As long as fast changes are found simultaneously in the time course of a significant number of marker tracks, they will be detected.

When data are filtered, knowledge of the character of the process underlying the signal is important. In many systems, the signal-to-noise ratio decreases with increasing frequency. In such a context, low-pass filtering in time is obvious. However, when tracking many markers that are part of a continuum, the degree of coherence in motion of the markers seems to be a valuable identifying property of the signal with respect to a noisy background, especially in the case of a signal with a relatively high-frequency content with respect to the 50-Hz video sampling rate. From this consideration, SVD is much more useful than low-pass filtering.

In the present example the SVD-filtering method is applied to two-dimensional marker tracks. However, there is no limitation to the number of dimensions. With minor modifications the method can be applied to a set of three-dimensional marker tracks. In the present example, the number of marker coordinates is generally larger than the number of samples in time. With a smaller number of markers or a longer observation period, the situation may be reversed. In that case, SVD should be applied after interchanging the role of time and space in the mathematical setup.

The SVD-filtering method enables a high resolution in detecting relative marker motion. When absolute positions are required, the degree of distortion of the camera imaging system has to be known precisely. Generally, a grid calibration in a single frame is not accurate enough to raise the precision of the calibration to the high level of accuracy reached in the multiframe experiments. Thus the improvement of positional accuracy by SVD filtering

forces us to pay more than usual attention to camera calibration.

In conclusion, the SVD filtering method presented here determines the positions of moving markers far more accurately than expected on the basis of limitations set by pixel resolution. The low level of positional noise is extremely useful in the determination of cardiac deformation, especially when high spatial resolution of gradients in deformation is required.

Address reprint requests to A. M. M. Muijtens, M. Sc., Dept. of Medical Information and Statistics, Univ. of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

Received 29 August 1988; accepted in final form 24 August 1989.

REFERENCES

1. BARD, Y. *Nonlinear Parameter Estimation*. New York: Academic 1974, p. 195.
2. DIXON, W. J. (Editor). *BMDP Statistical Software*. Berkeley: Univ. of California Press, 1985, p. 604.
3. GALLAGHER, K. P., R. A. GERREN, M. C. STIRLING, M. CHOY, R. C. DYSKO, S. P. MCANIMON, AND W. R. DUNHAM. The distribution of functional impairment across the lateral border of acutely ischemic myocardium. *Circ. Res.* 58: 570-583, 1986.
4. HATZE, H. The use of optimally regularized Fourier series for estimating higher-order derivatives of noisy biomechanical data. *J. Biomech.* 14: 13-18, 1981.
5. NUMERICAL ALGORITHMS GROUP Fortran Library Manual, Mark 9. Oxford, UK, 1982, vol. 4.
6. PEZZACK, J. C., R. W. NORMAN, AND D. A. WINTER. An assessment of derivative determining techniques used for motion analysis. *J. Biomech.* 10: 377-382, 1977.
7. PRINZEN, F. W., T. ARTS, T. T. PRINZEN, AND R. S. RENEMAN. Mapping of functional borderzone around ischemic myocardium. (Abstract). *Federation Proc.* 46: 1052, 1987.
8. PRINZEN, F. W., T. ARTS, G. J. VAN DER VUSSE, W. A. COUMANS, AND R. S. RENEMAN. Gradients in fiber shortening and metabolism across ischemic left ventricular wall. *Am. J. Physiol.* 250 (Heart Circ. Physiol. 19): H255-H264, 1986.
9. PRINZEN, T. T., T. ARTS, F. W. PRINZEN, AND R. S. RENEMAN. Mapping of epicardial deformation using a video processing technique. *J. Biomech.* 19: 263-273, 1986.
10. STEWART, G. W. *Introduction to Matrix Computations*. London: Academic 1973, p. 317-323.
11. WOOD, G. A., AND L. S. JENNINGS. On the use of spline functions for data smoothing. *J. Biomech.* 12: 477-479, 1979.

Extrapolation of incomplete marker tracks by lower rank approximation

Arno M.M. Muijtjens, Jef M.A. Roos, Theo Arts*, Arie Hasman, and Robert S. Reneman**

Departments of Medical Informatics and Statistics, Biophysics, and Physiology**,
Cardiovascular Research Institute Maastricht
University of Limburg, The Netherlands*

Published in Int. J. Biomed. Comput. 33: 219-239, 1993.

EXTRAPOLATION OF INCOMPLETE MARKER TRACKS BY LOWER RANK APPROXIMATION

ARNO M.M. MUIJTJENS^a, JEF M.A. ROOS^a, THEO ARTS^b, ARIE HASMAN^a
and ROBERT S. RENEMAN^c

Departments of ^aMedical Informatics and Statistics, ^bBiophysics and ^cPhysiology, Cardiovascular Research Institute Maastricht, University of Limburg (The Netherlands)

(Received December 28th, 1992)

(Accepted February 1st, 1993)

Motion and deformation of an object such as the heart may be measured by tracking optical or radio-paque markers. In the experimental situation markers may fail to be detected due to occlusion or lack of contrast. As a result a continuous marker track is observed in separated parts, which often cannot be directly identified as corresponding to one marker. This paper presents a method of extrapolating a partly known track by using information provided by the known track part and the available complete tracks of other markers. The extrapolations are obtained by iteratively fitting a lower rank matrix to the set of noisy, incomplete marker tracks. The performance is evaluated with computer-simulated data and data obtained in an animal experiment. In both cases 43% of the available complete tracks were made incomplete by removal of track parts varying in length from 3% up to 44%. For the simulated data comparison of the extrapolations with true signal values results in a root mean square (RMS) error about equal to the noise level. For the animal experiment, when comparing the extrapolations with the measured values, in images of 256×256 pixels, the RMS error was found to be ± 0.5 pixel, which is quite small relative to the total excursion of a marker (20 pixels). Estimation of the missing data by applying BMDPAM (BMDP Statistical Software Inc.) to the same data results in RMS errors which are about twice as high.

Key words: Motion; Algorithms; Signal processing; Image processing; Myocardium; Biomechanics

Introduction

In order to understand cardiac mechanics, regional motion and deformation of the cardiac wall has to be measured in physiological experiments. Existing techniques are based on two-dimensional echocardiography (2DE), magnetic resonance imaging (MRI) as well as following attached markers, visualized either optically or by the use of X-rays.

By the use of ultrasonic 2DE the motion of the heart can be measured non-invasively [1–3]. The resolution of the method is limited, and distinct identifiable points cannot be followed in three dimensions. Recent developments in MRI are of significant importance. By a new method called tagging, specific regions of the myo-

Correspondence to: A.M.M. Muijtjens, M.Sc., Dept. of Medical Informatics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

cardium are labeled and myocardial motion can be derived from the displacement and deformation of the tag patterns [4–8]. At the present state of the art, when using tagging in MRI, only the motion in the image plane direction is measured and the MR imaging procedure requires stable periodic motion during hundreds of periods. Nevertheless, the rapid developments in this field are very interesting and may lead to an MRI-based non-invasive method of following distinct tagged points in 3-D with sufficient resolution in space and time.

Regional motion and deformation may be measured by following markers attached to the heart in a sequence of images. In our laboratory the deformation of the heart was investigated in open thorax animal experiments by following about 50 optical markers which were attached to the surface of the left ventricle [9–12]. Figure 1 shows a resulting observed set of marker tracks.

In the coronary arterial tree, bifurcations can be visualized with cineangiography. They are used as natural markers when investigating heart motion in closed-chest experiments [13–16]. Radiopaque markers have often been implanted to serve as identifiable points in measuring heart motion, using single or biplane X-ray recordings of the heart [17–23]. Garrison et al. [19] developed a set-up to measure the 3-D motion of larger numbers of implanted radiopaque markers (up to 20). Manually indicated marker positions were digitized and stored on a disk to be analyzed by computer. Hannaford and Glantz [20] developed an algorithm to automatically track markers by the use of an adaptive linear predictor. This algorithm is employed

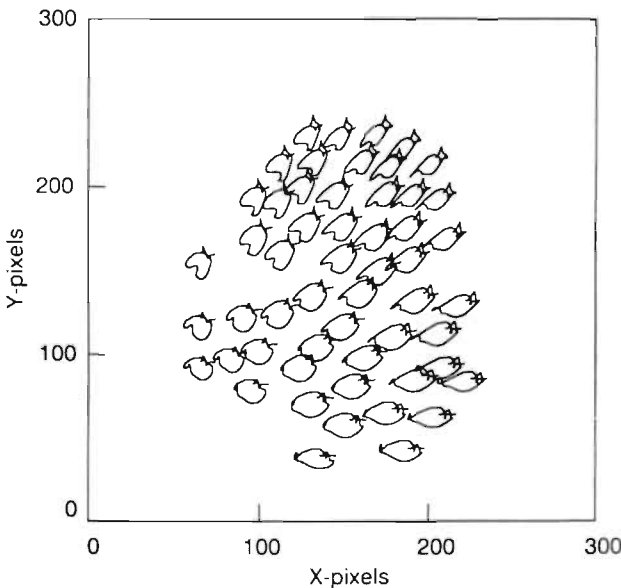


Fig. 1. Marker tracks during one cardiac cycle in an open-chest dog experiment. The motion through each loop is directed clockwise. The data are obtained in a 35×45 mm area of the anterior epicardial surface; 54 markers are tracked in 39 frames, resulting in an observation time of about 0.8 s.

by Niczyporuk and Miller [23] in their PC-based system which tracks up to 22 radiopaque markers.

At present, following the motion of implanted radiopaque markers by a biplane X-ray set-up seems to be most appropriate for obtaining accurate 3-D motion data of the heart with sufficient resolution in time and in space [5,18] in closed-chest dog experiments. Moreover, the set-up offers good opportunities for computerizing detection and tracking of videorecorded markers. Furthermore, the method is also appropriate for detecting deformation over much longer periods than one cardiac cycle, for instance for detecting gradual cardiac dilatation or myocardial growth.

Besides the detection of a marker, following a marker in time requires identification of the markers from one frame to the next. Identification of the markers is relatively simple if the number of markers is limited, their images are clearly distinct from the background, and they do not move too fast. In the experimental situation, however, these conditions are often not satisfied. For instance, when a marker image is missing in one frame of a sequence in time, tracking is interrupted. When the marker reappears, a new part of the track can be followed (Fig. 2). The problem is that after a gap, the newly started track often cannot be easily identified as an extension of a previously obtained track part. So, all parts of marker tracks should initially be treated as separate tracks and, as a result, the number of (incomplete) tracks is larger than the number of markers (Fig. 2, right panel).

Several techniques have been developed to extrapolate marker tracks. A straightforward method to extend tracks is to assume that the new position will be in a bounded region in the vicinity of the last position detected. If the velocity is too high, markers may move out of this region, resulting in interruption of the track. The markers may even move into regions where the extension of other markers is expected, resulting in falsely crossing-over of the tracks from one marker to another. Techniques have, therefore, been developed to obtain more accurate estimates of the new marker position. The track of a marker may be extended linearly on the basis

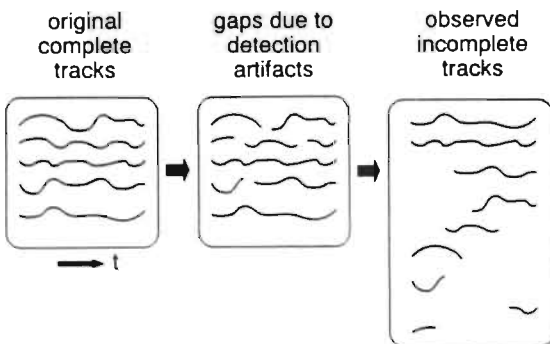


Fig. 2. Breaking up of marker tracks. Left panel: complete tracks. Middle panel: measured marker tracks (solid line) with missing parts (blanks). Right panel: observed set of marker track parts, ordered according to length in time. The problem is to reconstruct the tracks from the right to the left panel.

of the last two or more frames as if the velocity of the marker remained constant [19,24–27]. In the tracking procedure presented by Hannaford and Glantz [20] an adaptive linear predictor is applied. The new marker position is assumed to be a linear function of the last three positions in the track. The value of the coefficients is updated in each step of the tracking procedure such that the mean square prediction error is minimized. Occasionally, coherence of the motion of neighbouring markers may be used in predicting the extension of tracks. For example in measuring human body motion by following optical markers, the assumption of invariability of distance between markers on the same limb may be used [27]. In another study, in measuring epicardial deformation by following optical markers, track extensions have been estimated by assuming similarity of motion of neighbouring markers [9].

In the presence of detection errors, markers are observed in separate track parts. An accurate estimate of the extension of a track part is useful in finding the true continuation. This predicted extension may be used either to find the best matching continuation among available track parts, or to fill in the gaps between known parts of a marker track. The above-mentioned prediction methods may be used to estimate the position of a marker in a next frame, but the quality of the estimate rapidly deteriorates with the number of frames of the extension. The constraint that neighbouring markers have a similar motion cannot be used when markers moving in 3-D are projected in 2-D images. Closely projected markers may be distant in 3-D and, as a consequence, may have different motion patterns.

Earlier [28] it was shown that substantial noise reduction can be achieved in a marker position data matrix by using the sum of the first L major components of the Singular Value Decomposition as a signal estimate (SVD method). From the definition of the SVD [29], the procedure may be interpreted as the least squares fitting of a lower rank matrix to the marker position data. The accurate signal estimations obtained with the method indicate that a lower rank matrix is an adequate model for a set of marker tracks. However, the method requires a set of complete tracks without any missing data.

In the present study the SVD method has been modified so that the last requirement is not mandatory any more. In the new procedure a lower rank matrix is iteratively fitted to the available marker position data, including incomplete tracks (ILR method). The values obtained by the resulting lower rank matrix, at positions which correspond to missing data, are used as extrapolations of the incomplete marker tracks. In a computer simulation, the newly developed procedure is evaluated by comparing the resulting extrapolations with the known signal continuations. The method is also applied to tracks obtained in an animal experiment. Then, parts of the marker tracks are left out on purpose and the algorithm is used to reconstruct these parts. The reconstruction is compared with the actually measured continuation of the tracks. The performance of the ILR algorithm is compared with the results obtained with the program AM of the statistical package BMDP [30].

Lower Rank Model for Incomplete Marker Tracks

Markers attached to a moving or deforming object may be observed optically or with the help of X-ray equipment. In either case the marker movements are recorded

as a sequence of frames on cinefilm or videotape. The position of a marker in a frame is represented by the pair (x,y) , the image plane coordinates of the centroid of the markers projection. In order to measure the motion of a set of markers, their positions are detected and corresponding markers in successive frames are identified. The resulting tracks of markers in the image plane are represented by a set of chains of (x,y) coordinate pairs (Fig. 1). In general, some marker tracks are incomplete due to losing track. Hence, after primary detection and tracking a set of incomplete marker tracks is found (Fig. 2). The measured positions of markers associated with N tracks, observed over a time period of T frames is represented by an incompletely filled $2N \times T$ matrix \mathbf{V} :

$$\mathbf{V} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1T} \\ x_{21} & x_{22} & \cdots & x_{2T} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NT} \\ y_{11} & y_{12} & \cdots & y_{1T} \\ y_{21} & y_{22} & \cdots & y_{2T} \\ \vdots & \vdots & & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{NT} \end{bmatrix} \quad (1)$$

with $(x,y)_{ij}$ the position of the detected marker in track i at time frame j with $1 \leq i \leq N$ and $1 \leq j \leq T$.

Earlier a SVD-based lower rank model was presented for filtering a set of complete marker tracks [28]. For the set of incomplete tracks in matrix \mathbf{V} the corresponding model has been modified by assigning a weight to each element. The weight is equal to 0 in unknown parts of a track and is equal to 1 in detected parts. So, the new model may be written as:

$$\mathbf{W} \times \mathbf{V} = \mathbf{W} \times (\mathbf{C}\mathbf{B}^T + \mathbf{E}) \quad (2)$$

with \mathbf{W} a $2N \times T$ matrix of weights and $\mathbf{W} \times \mathbf{V}$ the element by element product of matrices \mathbf{W} and \mathbf{V} . \mathbf{C} is a $2N \times L$ matrix with orthogonal column vectors and \mathbf{B} a $T \times L$ matrix with orthogonal column vectors of unit length. The matrix product $\mathbf{C}\mathbf{B}^T$ represents the true signal without noise, which is assumed to have a rank equal to $L < \text{minimum}(2N, T)$. This definition of the true signal corresponds to reconstructing a signal matrix by using only the first L components of its SVD, which is analogous to the model used in SVD filtering a complete matrix as presented earlier [28]. Within the framework of this analogy, the columns of the matrices \mathbf{C} and \mathbf{B} correspond to the set of first L left and right eigenvectors of the SVD of a complete matrix \mathbf{V} , with the length of the k th column vector of matrices \mathbf{B} and \mathbf{C} equal to 1 and equal to the k th singular value, respectively [29]. The difference between the true signal and the actually measured signal is represented by a $2N \times T$ matrix \mathbf{E} of errors.

The elements of matrices \mathbf{C} and \mathbf{B} are fitted to the observed track parts such that

for a given rank L the sum of the squared elements of \mathbf{E} , as defined in Eqn. (3), is minimum.

$$\psi(\mathbf{C}, \mathbf{B}) = \sum_{i=1}^{2N} \sum_{j=1}^T e_{ij}^2 = \sum_{i=1}^{2N} \sum_{j=1}^T w_{ij} \left(v_{ij} - \sum_{k=1}^L c_{ik} b_{jk} \right)^2 \quad (3)$$

In criterion (3) the variables e_{ij} , w_{ij} , v_{ij} , c_{ik} and b_{jk} are elements of matrices \mathbf{E} , \mathbf{W} , \mathbf{V} , \mathbf{C} and \mathbf{B} , respectively. In searching the minimum of ψ a necessary condition is a zero gradient of ψ with respect to all unknown elements of matrices \mathbf{C} and \mathbf{B} . This yields (see Appendix A) the next two sets of equations:

$$\begin{bmatrix} \sum_{j=1}^T w_{ij} b_{j1} b_{j1} & \cdots & \sum_{j=1}^T w_{ij} b_{j1} b_{jL} \\ \vdots & & \vdots \\ \sum_{j=1}^T w_{ij} b_{j1} b_{jL} & \cdots & \sum_{j=1}^T w_{ij} b_{jL} b_{jL} \end{bmatrix} \begin{bmatrix} c_{i1} \\ \vdots \\ c_{iL} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^T w_{ij} b_{j1} v_{ij} \\ \vdots \\ \sum_{j=1}^T w_{ij} b_{jL} v_{ij} \end{bmatrix} \quad i = 1, 2, \dots, 2N \quad (4)$$

$$\begin{bmatrix} \sum_{i=1}^{2N} w_{ij} c_{i1} c_{i1} & \cdots & \sum_{i=1}^{2N} w_{ij} c_{i1} c_{iL} \\ \vdots & & \vdots \\ \sum_{i=1}^{2N} w_{ij} c_{i1} c_{iL} & \cdots & \sum_{i=1}^{2N} w_{ij} c_{iL} c_{iL} \end{bmatrix} \begin{bmatrix} b_{j1} \\ \vdots \\ b_{jL} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{2N} w_{ij} c_{i1} v_{ij} \\ \vdots \\ \sum_{i=1}^{2N} w_{ij} c_{iL} v_{ij} \end{bmatrix} \quad j = 1, 2, \dots, T \quad (5)$$

The elements w_{ij} and v_{ij} are known and the elements b_{ij} and c_{ij} have to be calculated. Equations (4) and (5) are solved iteratively. Substituting the elements b_{ij} of an initial guess $\mathbf{B}^{(1)}$, an estimate $\mathbf{C}^{(1)}$ of \mathbf{C} is found by solving Eqn. (4). An improved estimate $\mathbf{B}^{(2)}$ results from solving Eqn. (5) having substituted the elements c_{ij} of $\mathbf{C}^{(1)}$. Generally this procedure of repeated solving of Eqns. (4) and (5) converges to a unique solution [31]. The iteration is stopped if

$$1 - \sum_{k=1}^L \frac{|\mathbf{b}_k^{(i)} \cdot \mathbf{b}_k^{(i-1)}|}{|\mathbf{b}_k^{(i)}| |\mathbf{b}_k^{(i-1)}|} < \delta_1 \quad (6)$$

The dot represents the inner vector product and $\mathbf{b}_k^{(i)}$ is the k th column of \mathbf{B} as estimated in the i th iteration. We have chosen $\delta_1 = 10^{-8}$.

Extrapolation by Iterative Lower Rank (ILR) Approximation

In the ILR algorithm the model \mathbf{CB}^T in Eqn. (2) is estimated stepwise, starting with a matrix of rank $L = 1$. In each next step the rank is increased by adding a column to matrices \mathbf{C} and \mathbf{B} . With increasing rank, the energy contribution of the new term to the model \mathbf{CB}^T decreases. Rank increase is stopped when the energy of additional contributions is of the order of the estimated noise level. The stop criterion is similar to the one used in SVD filtering [28]:

$$\frac{\sum_{i=1}^{2N} \sum_{j=1}^T w_{ij} \left(v_{ij} - \sum_{k=1}^L \hat{c}_{ik} \hat{b}_{jk} \right)^2}{[2N \cdot T - S - (2N + T - L)L]\sigma^2} < \delta_2 \quad (7)$$

where \hat{c}_{ik} and \hat{b}_{jk} are estimates of the elements of matrices \mathbf{C} and \mathbf{B} . Elements w_{ij} and v_{ij} are defined in relation to Eqn. (3). The nominator indicates the product of σ^2 and the number of degrees of freedom of the residual sum of squares, where σ^2 represents the variance of the noise in each element of \mathbf{V} and S is equal to the number of missing elements in \mathbf{V} . We have chosen $\delta_2 = 1.2$.

Often the stop criterion presented in Eqn. (7) is adequate; however, in some cases, when the rank is increased, the new solution according to Eqns. (4) and (5) appears to be quite different in some of the lower rank terms. This indicates that collinearity is introduced in the fit and as a result the accuracy of the extrapolations decreases. When this situation occurs, the procedure is also stopped. Extrapolation accuracy is indicated to deteriorate when the length of the column vectors of \mathbf{C} in iteration L deviates too much from the corresponding values in iteration $L-1$:

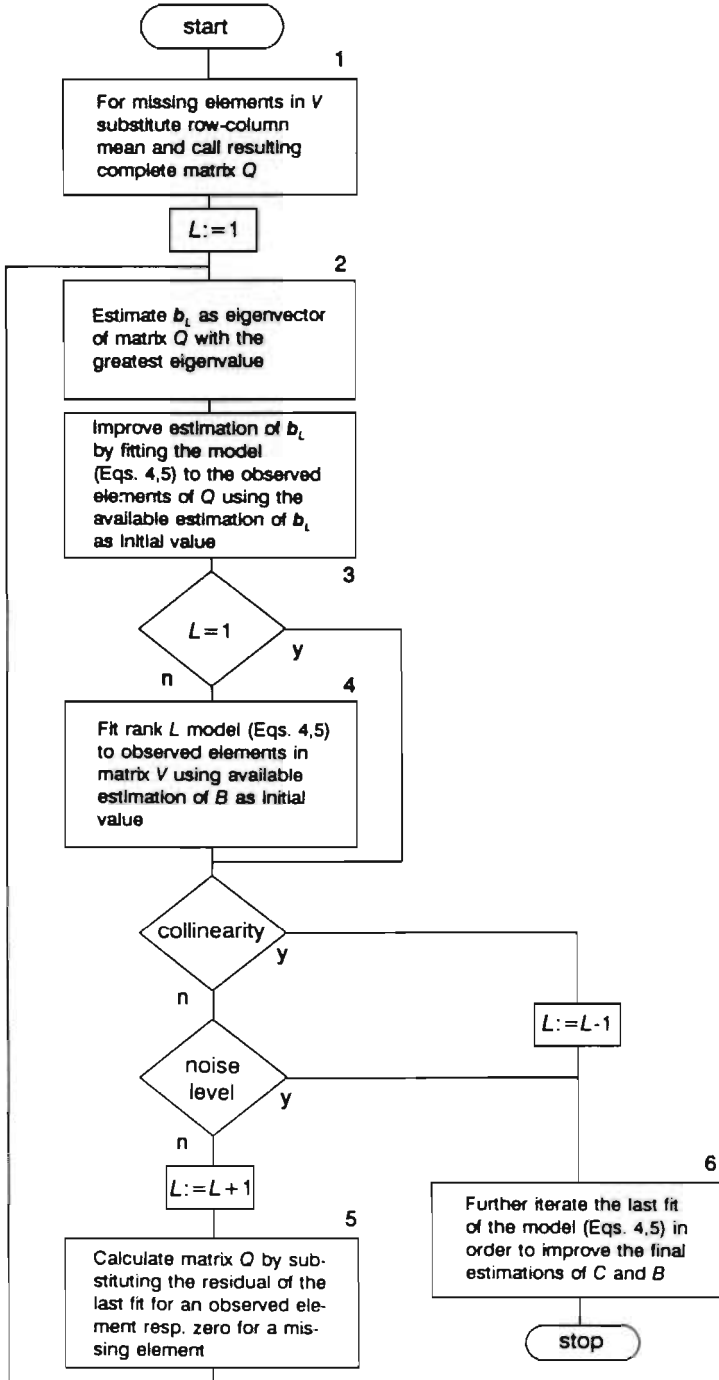
$$\sum_{k=1}^{L-1} \frac{|\gamma_k^{(L)} - \gamma_k^{(L-1)}|}{\gamma_k^{(L-1)}} > \delta_3 \quad (8)$$

Parameter $\gamma_k^{(L)}$ represents the squared length of the k th column vector of \mathbf{C} as found in the L th step of the algorithm and the threshold is set to $\delta_3 = 0.15$. If criterion (8) is satisfied, the L th iteration is ignored and the model in iteration $L-1$ is considered as the final solution.

The flow chart in Fig. 3 shows the structure of the algorithm. Starting with $L = 1$, the rank L of the model \mathbf{CB}^T is iteratively incremented by 1, i.e. the matrices \mathbf{C} and \mathbf{B} are extended with one column and fitted to the observed part of matrix \mathbf{V} . At the start of the algorithm (module 1), matrix \mathbf{Q} is filled with the observed elements of matrix \mathbf{V} , supplemented by a weighted sum of row and column averages (Eqn. 9) for each element v_{ij} which is missing.

$$Q_{ij} = \frac{s_j^2 \bar{v}_i + s_i^2 \bar{v}_j}{s_i^2 + s_j^2} \text{ when } v_{ij} \text{ is missing} \quad (9)$$

In this expression \bar{v}_i and \bar{v}_j are equal to the averages of the observed elements in the i th row and the j th column of matrix \mathbf{V} , respectively, and s_i^2 and s_j^2 are the corresponding variances.



In module 2 an initial value of \mathbf{b}_1 , the first column of matrix \mathbf{B} , is obtained by applying standard SVD procedures to the complete matrix \mathbf{Q} , which results from module 1. The estimation of \mathbf{b}_1 is improved in module 3 by fitting a rank 1 model to the observed elements of \mathbf{Q} by repeatedly solving Eqns. (4) and (5). For $L = 1$ this completes the calculations and module 4 is skipped. Next the collinearity (8) and the noise level stop criterium (7) are checked, and if neither of the two is satisfied the rank of the model is incremented by 1.

The second iteration starts by filling matrix \mathbf{Q} with the current residuals of the observed elements in \mathbf{V} , supplemented by zero values for the missing elements (module 5). On the basis of matrix \mathbf{Q} an initial value for \mathbf{b}_2 is calculated in module 2 and an improved estimate is obtained in module 3. The resulting estimate for \mathbf{b}_2 and the estimate of \mathbf{b}_1 , which has been obtained in the first iteration, are used as initial values in the actual fit of the rank 2 model to the observed elements of matrix \mathbf{V} (module 4). Hence, in each iteration all parameters of the extended model are re-estimated by repeated solving of Eqns. (4) and (5) in module 4. In this calculation, the estimations obtained in a former iteration are used as initial value.

Iteration proceeds until one of the conditions (7) or (8) is satisfied. At the end of the algorithm the fit is further improved by iteratively solving (4) and (5) until condition (6) with $\delta_1 = 10^{-8}$ is satisfied. For the iterations in module 4 a less stringent value of $\delta_1 = 10^{-5}$ is used in order to save computing time.

The ILR algorithm is implemented in *FORTRAN* on a VAX 750 computer. Several subroutines are obtained from the Numerical Algorithms Group (NAG) *FORTRAN* library [32]. Routines F03AEF and F04AGF are used to solve the system of L linear equations with L unknowns represented by Eqns. (4) and (5). In module 2 of the algorithm (Fig. 3) the eigenvector corresponding to the greatest eigenvalue is calculated by using routines F01AGF, F02BFF, F02SDF and F01AHF. On the VAX 750 the run-time of the ILR algorithm is equal to about 10 min for a 100×50 matrix with 10% of the data missing.

Methods

The performance of the algorithm is evaluated (1) for computer-simulated data and (2) for data obtained in an animal experiment. In both cases the original data consist of a set of complete marker tracks. To evaluate the extrapolation method, several tracks were partly removed. The removed parts were reconstructed by extrapolation of the set of observed track parts by using the ILR method. The method was assessed by comparing the extrapolations with the known removed track parts.

Fig. 3. Flowchart of the ILR algorithm. An SVD-based lower rank model \mathbf{CB}^T is fitted iteratively to the incomplete marker tracks in matrix \mathbf{V} . Starting with a model of rank $L = 1$, the rank of the fitted model is iteratively incremented by 1, i.e. matrices \mathbf{C} and \mathbf{B} are extended with one column. Rank increasing proceeds until one of the stop criteria is reached. A first estimate of \mathbf{b}_L , the new column of matrix \mathbf{B} , is obtained by calculating the eigenvector associated with the greatest eigenvalue of the residual matrix \mathbf{Q} . While iterating, this matrix contains the residuals of the former fit supplemented with zeros for the missing elements. At the start of the algorithm, matrix \mathbf{Q} is filled with the elements of \mathbf{V} , supplemented with a weighted row and column mean for the missing elements.

Computer generation of data on marker motion

In the computer simulation a 100×50 matrix was generated to describe the motion of 50 markers sampled at 50 moments in time. The x - and y -coordinates correspond to the first and second 50 rows, respectively. The coordinate data contained the signal to which noise was added. The signal was obtained by generating two independent sets of 40 orthogonal vectors being used to fill the columns of matrices C and B , respectively. Thus, the rank of the resulting matrix CB^T was equal to 40. Initially the elements of each column vector were chosen to be random samples from a zero-mean gaussian distribution. Starting from the first randomly generated vectors each subsequent vector is obtained by removing the projection of the next random vector on the space spanned by all former vectors. The length of each vector was set to unity by multiplication with a proper scalar. Then the column vectors in C were multiplied with the singular values, the distribution of which is based on physiological experiments. The logarithm of the first 12 singular values found in the SVD of an animal experiment [28] was approximated by a linear function of the rank. Finally, generating simulated data was completed by adding uncorrelated gaussian noise to each element of the signal matrix CB^T with a signal to noise energy ratio of about 10^3 .

Acquisition of experimental data on marker motion

The motion of about 50 markers was obtained in a physiological experiment. In an open-chest canine preparation the heart was exposed and 54 white markers were attached to the epicardium [9] (Fig. 4). Their motion was observed with a video camera and recorded on videotape. After the experiment the recorded video signal was digitized by a 4-bit A/D-converter at a sampling rate of 5 MHz. Up to 64 frames of 256×256 pixels, separated 20 ms in time, could be stored in a 2-Mb Digital Video Memory (DVM, designed and built in our laboratory). The video data in the DVM were accessed by a personal computer (Laser AT) for image analysis. Marker tracks are represented as arrays of (x, y) coordinates sampled in time each 20 ms (Fig. 1). In the used set-up all observed marker tracks were complete. The 54 markers were followed in 39 frames resulting in a 108×39 matrix V .

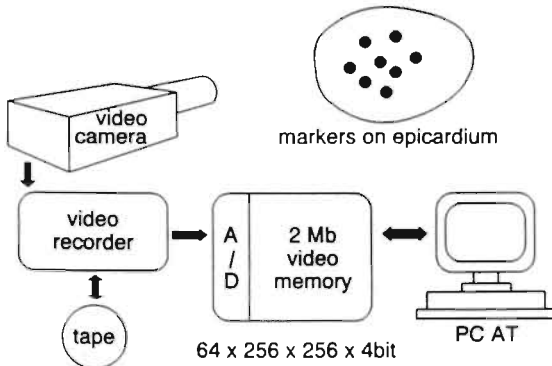


Fig. 4. Block diagram of the image acquisition system which is used to measure epicardial motion in open-chest experiments on dogs. The method is based on measuring the motion of optical markers by the use of video techniques.

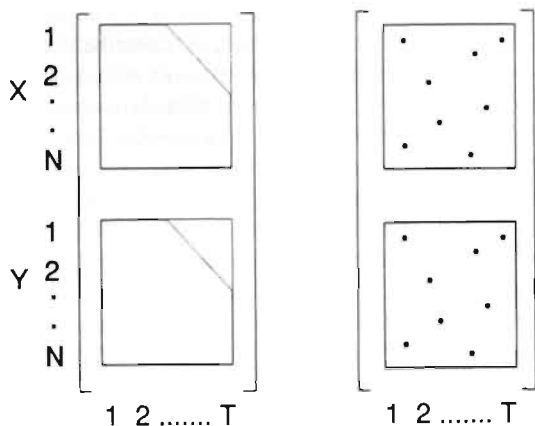


Fig. 5. The two missing data structures used in the evaluation of the performance of the extrapolation method. The shaded parts are removed for the analysis. At the left a triangular part of the set of tracks is removed and at the right positions are removed at random.

Evaluation of the Iterative Lower Rank extrapolation method

In the evaluation two missing data structures were used. To evaluate extrapolation over relatively long time intervals, a triangular part of the coordinate data matrix was removed before analysis (Fig. 5, left panel). To evaluate the quality of interpolation over gaps in the tracks, data were left out at random in the set of tracks, while the relation between separate track parts was retained (Fig. 5, right panel). Note that the distinction between interpolation and extrapolation is not essential. In the rest of the paper the term extrapolation will be used.

In the case of computer-simulated data, the extrapolation error is defined as the difference between the extrapolation and the true signal without noise. In the animal experiment, however, the signal could not be separated from noise, so the extrapolation error is assessed by comparing the extrapolations with the actually measured values. To compare the extrapolation results of the animal experiment with those of the computer-simulated data, the extrapolation error was also evaluated by calculating the difference between the extrapolation and the raw data, including the noise. Errors were quantified by the root mean square (RMS) value. The distribution of the absolute value of the errors was calculated in terms of the 50%, 75% and 93.75% percentiles. These percentiles refer to the central location, the spread and the outliers of the distribution, respectively. The error level is presented in units of σ , the RMS value of the noise. For the computer-simulated marker data the value of σ is exactly known. In the case of the experimental data the noise level found in Ref. 28, 0.007 pixel^2 , was used as an estimate of σ^2 .

For the triangular and the random missing data structure, ten repeated tests were performed for amounts of missing data equal to 10%, 20%, 25%, 40%, 50% and 75%. For every percentage of missing data the RMS value of the extrapolation errors and the percentiles of the error distribution were calculated. In the case of computer-simulated data the 100×50 matrix was generated ten times, each time with a dif-

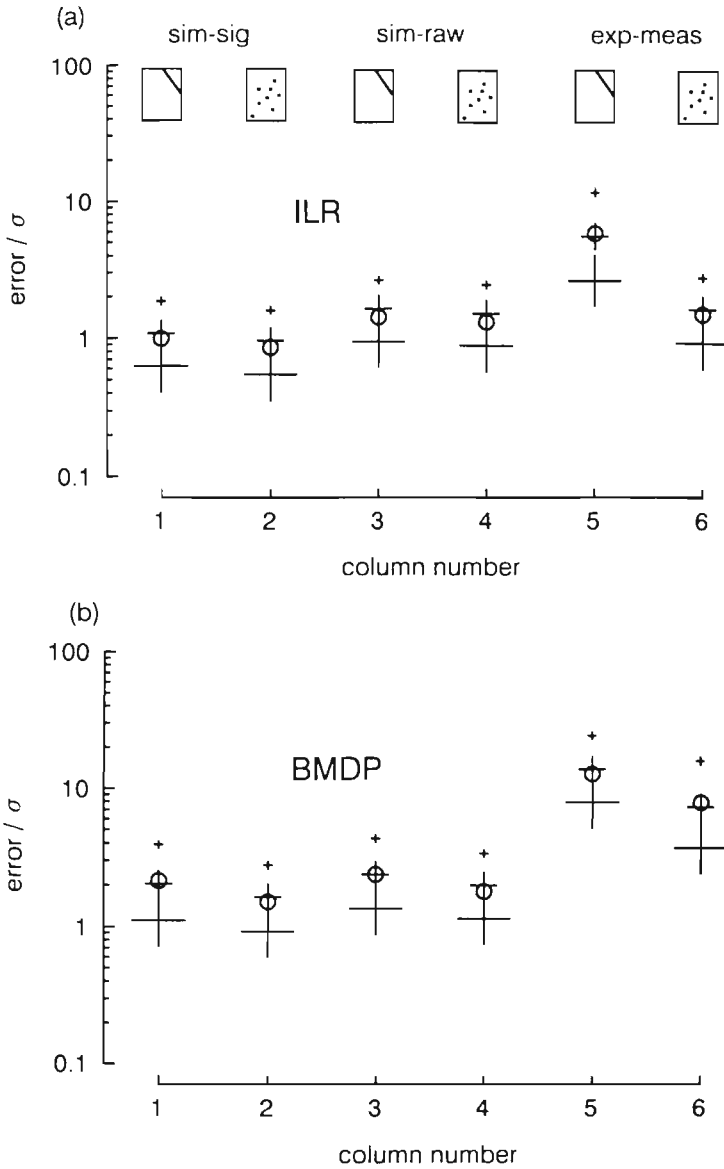


Fig. 6. Distribution and RMS value of the extrapolation error in tests with 10% missing data. The error is presented in units of σ , the RMS value of the noise. Each distribution summarizes the errors found in ten repeated tests. Crosses of decreasing size indicate the 50%, 75% and 93.75% percentiles of the distribution of the absolute value of the error. The RMS value is indicated in the plot with a circle. Panels a and b show the results for ILR extrapolation and for the program BMDPAM, respectively. The first two columns present the results obtained with computer simulated data after removal of a triangular part and after removal of data at random, respectively. The error is calculated by using the true signal value as a reference. Columns three and four show the same results but now the reference is the signal including the noise (raw data). The last two columns represent the results for similar tests based on experimental data. The measured signal is used as a reference in calculating the extrapolation error.

ferent random noise sequence as well as a different set of random signal vectors. For the experimental data only one set of complete data was used. Before the upper right triangle of the matrix was removed the tracks were permuted randomly. In case of random removal of data, the missing x, y -element pairs in each test were selected randomly in the data matrix.

The results of the ILR algorithm were compared with those of the program AM in the statistical package BMDP [30] which estimates missing data by linear regression based on maximum likelihood estimates of the correlation matrix [33].

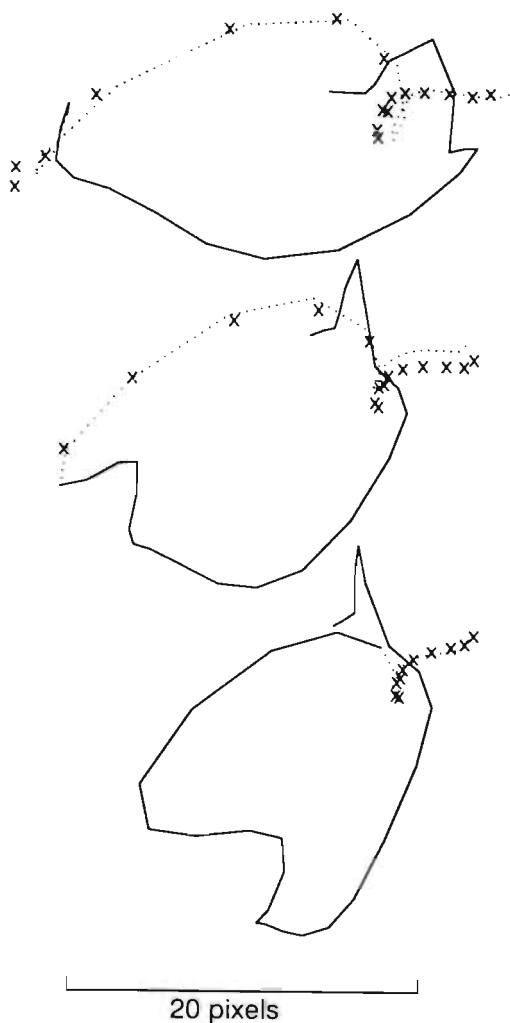


Fig. 7. Extrapolation of partly observed tracks in a test based on experimental data with a 10% triangular part removed. Three tracks with 44%, 38% and 26% missing positions are shown. The observed part of a track is indicated by a solid line, crosses indicate positions found by ILR extrapolation and dotted lines indicate the part of the measured track which was removed in the test.

The evaluation was started with testing the performance of the method with 10% of the data removed. Then the percentage missing data was increased to assess the limits of performance. For the tests with 10% missing data the extrapolation error was investigated as a function of the rank of the fitted model. The resulting curves are used to assess the ability of the ILR method to find the rank with minimum extrapolation error.

Results

In Fig. 6 the distribution of the absolute value of the calculated extrapolation error is shown for the tests with 10% of the data removed. In panel a the results obtained with the ILR algorithm are shown after removal of a triangular part (columns 1, 3 and 5) as well as after random data removal (columns 2, 4 and 6). In panel b corresponding results are presented when estimating the missing data by using BMDPAM. In the first two columns the results for the computer simulations were compared with the true signal as a reference, in columns 3 and 4 the results were compared with the raw data, including noise. In columns 5 and 6 the results are shown for the animal experiment, while using the measured value as a reference in calculating the error. The distribution of the absolute value of the extrapolation error is represented by the percentiles 50%, 75% and 93.75%, which are indicated by crosses in decreasing order of size. In Fig. 6 the RMS value of the error is indicated by a circle.

To illustrate the accuracy of the ILR algorithm in the animal experiment after removal of a 10% triangular part of the data, the extrapolation of three experimentally measured tracks is shown in Fig. 7. The available parts of the tracks are indicated by a solid line. The dotted lines represent the removed parts with time fractions of 44%, 38% and 26% of the complete tracks, respectively. The crosses indicate the estimated positions which were obtained with the ILR method.

In order to assess the limitations of the method, the RMS extrapolation error was investigated as a function of the percentage of missing data (Fig. 8). The results of tests based on computer-simulated data and tests based on data obtained in an animal experiment are presented in panels a and b, respectively. Each panel shows the results obtained with the ILR algorithm (solid line) and with the program BMDPAM (broken line) for removal of a triangular part (triangle) and for random data removal (circle). Each symbol in a curve corresponds to the RMS error found in ten repeated tests.

The performance of the stop criteria with respect to terminating the algorithm at an optimal rank is illustrated in Fig. 9 for computer-simulated data (panel a) and experimental data (panel b). The results after removal of a 10% triangular part are shown in the left part of each panel and the results after 10% random data removal are shown in the right part. The plotted bundle of curves represents ten repeated tests, in which the RMS extrapolation error is shown as a function of the rank of the fitted model. In each curve the final point of the fit is presented. A circle indicates termination by reaching the noise level (7) and a square termination by the collinearity criterion (8). A vertical dashed line indicates the optimal rank which was found by applying the SVD method for complete tracks [28] to the complete data.

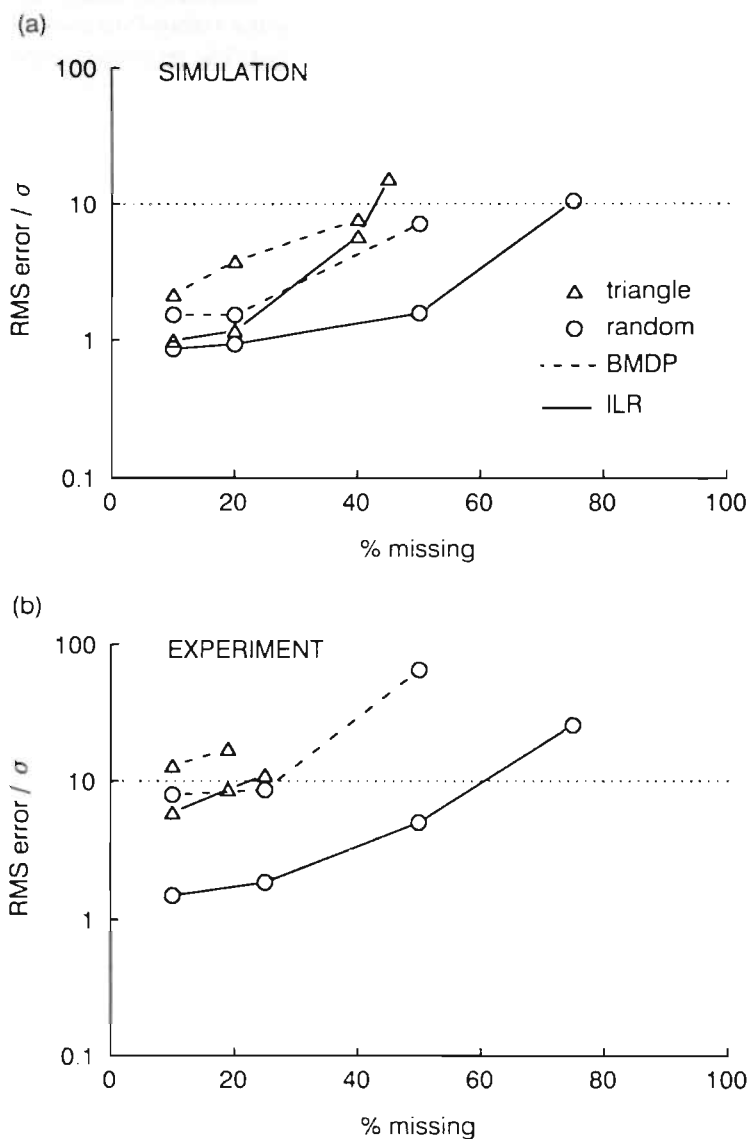


Fig. 8. The RMS extrapolation error as a function of the percentage missing data. The error is presented in units of σ , the RMS value of the noise. Results of tests based on computer-simulated data (panel a) and on experimental data (panel b) are shown for ILR (solid line) and BMDPAM (broken line) and for, respectively, a triangular part removed (triangle) and data removed at random (circle). Each point represents the RMS error found in ten repeated tests. The 10σ level is indicated by a horizontal dashed line.

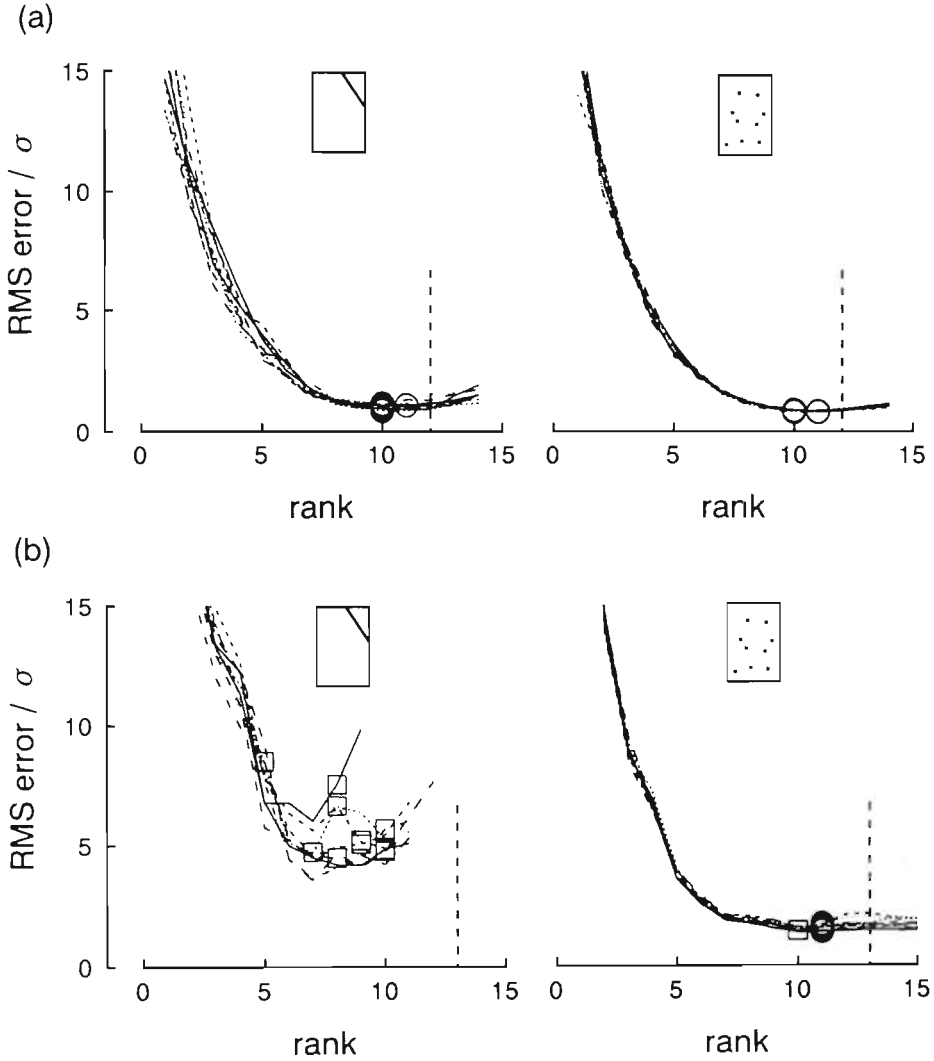


Fig. 9. The RMS extrapolation error as a function of the rank of the fitted model. The error is presented in units of σ , the RMS value of the noise. Each curve corresponds to the results obtained in one of ten repeated tests. Results are shown for computer-simulated data and experimental data (panels a and b) with a 10% part removed. The error is calculated by using as a reference the true signal value in the case of the simulations and the measured value with the experimental data. Presented are the results with a triangular part removed (left) and with positions removed at random (right). The final point of the fit and the reached stop criterion are indicated: a circle corresponds to the noise level criterion (7) and a square to the collinearity criterion (8). A vertical broken line indicates the optimal rank which is found when the SVD method for complete tracks is applied to the complete data.

Discussion

In this study the Iterative Lower Rank approximation algorithm (ILR algorithm) to extrapolate incomplete marker tracks over relatively long time intervals is developed and evaluated. The ILR algorithm is a modification of the SVD filtering method described earlier by us [28], which can only be used for a complete set of data. The performance of the ILR algorithm was evaluated in tests with computer-generated data as well as in a practical situation on the basis of data obtained in an animal experiment.

For experimental data with a 10% triangular part missing, ILR extrapolation results in a RMS extrapolation error and a 93.75 percentile value of 6σ and 12σ , respectively (Fig. 6, panel a, column 5). If under similar conditions data are missing at random (Fig. 6, panel a, column 6) the performance is better by a factor of 4. Although not as extreme, a similar behavior is found for the computer-simulated data (panel a, columns 3 and 4). The relatively high error level found for the experimental data with a triangular part missing is related to an early occurrence of colinearity in the fit (Fig. 9, left part of panel b).

Comparing the two panels of Fig. 6, ILR extrapolation performs substantially better than BMDPAM with respect to the resulting RMS level of the error as well as with respect to the occurrence of outliers. The RMS value and the percentiles of the extrapolation error obtained with the ILR algorithm (panel a) are about 50% lower than the levels obtained with the program BMDPAM (panel b). Regression-based procedures, like the one implemented in BMDPAM, estimate the missing data by the correlation in only one direction of the data matrix. Usually, the columns of the data matrix are assumed to represent correlated variables and predictions are based on the information in the correlation matrix of these variables. In the SVD-based lower rank model, the columns of matrices **B** and **C** are the eigenvectors of matrices $\mathbf{V}^T\mathbf{V}$ and $\mathbf{V}\mathbf{V}^T$, respectively. These matrices represent the column as well as the row correlations of the elements in the data matrix **V**. So the ILR extrapolations are based on correlations in both directions of rows and columns. This may explain the better performance of the SVD-based method if compared with the results of the regression-based method implemented in BMDPAM.

In a practical situation, as simulated on the basis of the experimental data with a 10% triangular part removed, the RMS error of the ILR extrapolation is found to be 6σ . The noise variance, σ^2 , for these data equals 0.007 pixel^2 [28]. Hence, the RMS extrapolation error is equal to 0.5 pixel. For comparison, the total excursion of a marker is about 20 pixels. The resulting extrapolated positions are quite accurate as is illustrated in Fig. 7 by the close match of the extrapolations (crosses) and the earlier removed track parts (dashed lines).

Assuming a maximum acceptable level of the RMS extrapolation error of 10σ (Fig. 8, dotted horizontal line), the maximum percentages missing data for ILR extrapolation with computer-simulated data (Fig. 8, panel a, solid lines) were found to be equal to 42% and 75% for a triangular part missing (triangle) and for randomly missing data (circle), respectively. For the tests based on experimental data, these percentages are found to be 22% and 60%, respectively. Note that missing a triangle of 22% corresponds to a missing percentage per row ranging from 5 to 65%. Similarly, 42% missing corresponds with a range of 6% to 90% missing per row. Comparing

the ILR results with the results obtained with BMDPAM, it is found that the BMDPAM curves reach the threshold error value at a lower percentage of missing data (Fig. 8).

The results on the RMS extrapolation error as a function of the rank of the fitted model (Fig. 9) are similar for the computer simulated data (panel a) and the experimental data with data removed at random (panel b, right part). In the majority of the tests, iteration proceeded until the noise level criterion (7) was reached at ranks 10 or 11 and the resulting extrapolation error has an RMS value between 1.0σ and 1.5σ . For the experimental data with a triangular part missing (panel b, left part), however, iteration was terminated by the collinearity criterion (8) at ranks in the range of 5 to 10. As a result, in this case the level of the error was higher, being approximately 6σ . However, when using the criteria as mentioned, under all circumstances investigated the extrapolation error was close to the minimum value, indicating that the set of stop criteria has been adequately chosen.

In the ILR algorithm (Fig. 3) the data in matrix \mathbf{Q} are used to calculate an initial estimate of the new column of matrix \mathbf{B} by standard SVD procedures. While iterating matrix \mathbf{Q} is filled with the residuals of the fit of the model to matrix \mathbf{V} in the former step of the algorithm, supplemented with zero values for missing elements. At the start of the algorithm (Fig. 3 module 1) the missing values in \mathbf{Q} are replaced by the weighted column and row mean defined in Eqn. (9) instead of zero values, because in our set-up the signals contain a substantial DC component.

Iterative solving of Eqns. (4) and (5) is terminated by criterion (6). In this criterion the correlation between successive estimations of the parameters is used to quantify stationarity of the parameter values. For the threshold value we have chosen $\delta_1 = 10^{-8}$. Thus three to five iterations were needed at lower ranks and 10 to 20 at higher ranks of the model. To save computing time, during incrementing the rank of the model, the threshold is set to $\delta_1 = 10^{-5}$.

The increase of the rank of the model \mathbf{CB}^T is terminated if the residuals of the fit reach the expected noise level. This is indicated by criterion (7), which is similar to the criterion used in the SVD filtering method [28]. The threshold is chosen to be slightly higher than unity ($\delta_2 = 1.2$) to avoid convergence problems. For criterion (7) the variance of the noise should be known. When using experimental data, often this variance has to be estimated empirically [28]. When using computer-generated data σ^2 is exactly known.

Besides (7), a second criterion was useful. In some situations the extrapolation error appeared to increase at ranks far below the rank where the residual energy reaches the level of the noise. The increase of errors occurs together with substantial changes in the length of the column vectors of \mathbf{C} in successive steps of the algorithm. Interpreting the columns of matrix \mathbf{B} as basic time functions, each i th marker coordinate as a function of time is composed of these basic functions. Using the complete time interval, the basic functions in \mathbf{B} are orthogonal by definition. However, if part of the time interval cannot be taken into account due to missing data, the functions may be correlated significantly on the remaining interval. This collinearity may cause substantial inaccuracy in the estimations of the corresponding coefficients in the matrix \mathbf{C} and, as a result, increasing extrapolation errors. Collinearity is more likely to occur when the basic functions are relatively smooth and the missing points are concentrated in a specific time interval. This may account for the frequent occur-

rence of collinearity at relatively low ranks in the tests of the experimental data with a triangular part missing (Fig. 9, left part of panel b). The changes in the length of the column vectors of \mathbf{C} as found in successive iterations are used in the algorithm to indicate increasing extrapolation errors. The corresponding collinearity stop criterion is defined in Eqn. (8). A threshold level of $\delta_3 = 0.15$ is found to be appropriate, as is indicated by the position of the squares in Fig. 9 which are all at, or close to, the minimum extrapolation error.

The results presented in this study are obtained by applying the ILR algorithm to a matrix containing both the x and y coordinate data. The algorithm, however, could also be applied to the x and y coordinate data matrices separately. The performance of the ILR algorithm has also been tested for this situation. For the experimental data after removal of a 10% triangular part, the tests resulted in an RMS extrapolation error equal to 3σ , hence, an improvement of 50% compared with the result of combined treatment of x and y . In a test for the same data and 10% missing at random no significant change was found in the error level. Tests based on the computer-simulated data resulted in a slightly increased extrapolation error when x and y are treated separately. Hence, it is not clear whether separate treatment of x and y is advantageous.

Conclusions

(1) Application of the ILR method to incomplete marker tracks results in accurate extrapolations which are appropriate to be used in the rejoining of parts of broken marker tracks. In tests based on data obtained in an animal experiment, with a 10% triangular part removed in 54 tracks observed in 39 frames, an RMS extrapolation error of 6σ results, σ being equal to the RMS value of the noise. This corresponds with an error of 0.5 pixels, while the total excursion of a marker is about 20 pixels.

(2) Application of the ILR method to marker track data with data removed at random results in accurate estimations of the missing data. With 10% missing data, tests based on experimental data as well as tests based on computer-simulated data result in an RMS error of about 1.5σ .

(3) The use of the program BMDPAM results in RMS error levels which are about twice as high as the levels obtained with the ILR method.

(4) Assuming the maximum acceptable level of the extrapolation error to be equal to 10σ , for experimental data with a triangular part missing, the maximum percentage of missing data is 22% when using ILR extrapolation. With data missing at random the maximum is at 60%. When using BMDPAM these percentages are found to be lower: less than 10% and 25%, respectively.

Appendix A

Criterion (3) is a function of the elements of the unknown matrices \mathbf{C} and \mathbf{B} . For the minimum of this function it holds

$$\frac{\partial \psi}{\partial c_{qr}} = 0 \quad q = 1, 2, \dots, 2N; \quad r = 1, 2, \dots, L \quad (10)$$

$$\frac{\partial \psi}{\partial b_{im}} = 0 \quad t = 1, 2, \dots, T; m = 1, 2, \dots, L \quad (11)$$

Condition (10) applied to (3) results in

$$\sum_{i=1}^{2N} \sum_{j=1}^T w_{ij} \cdot 2 \cdot \left(v_{ij} - \sum_{k=1}^L c_{ik} b_{jk} \right) \left(- \sum_{s=1}^L \frac{\partial c_{is}}{\partial c_{qr}} b_{js} \right) = 0 \quad q = 1, 2, \dots, 2N; \\ r = 1, 2, \dots, L \quad (12)$$

Since the last factor in (12) is equal to 0 for $(i, s) \neq (q, r)$, we have

$$\sum_{j=1}^T w_{qj} \left(v_{qj} - \sum_{k=1}^L c_{qk} b_{jk} \right) b_{jr} = 0 \quad q = 1, 2, \dots, 2N; r = 1, 2, \dots, L \quad (13)$$

Similarly, applying condition (11) to (3) results in

$$\sum_{i=1}^{2N} w_{it} \left(v_{it} - \sum_{k=1}^L c_{ik} b_{tk} \right) c_{im} = 0 \quad t = 1, 2, \dots, T; m = 1, 2, \dots, L \quad (14)$$

Equations (13) and (14) can be rewritten to Eqns. (4) and (5), respectively.

References

- 1 Mailloux GE, Langlois FL, Simard PY and Bertrand M: Analysis of heart motion from two-dimensional echocardiograms by velocity field decomposition. In *Proc IEEE Conf Computers in Cardiology*, Leuven, Belgium, 1987, pp. 441–444.
- 2 Mailloux GE, Bleau A, Bertrand M and Petitclerc R: Computer analysis of heart motion from two-dimensional echocardiograms. *IEEE Trans Biomed Eng.* 34 (1987) 356–364.
- 3 Meunier J, Bertrand M and Mailloux G: Speckle tracking in echocardiography of the myocardium: a model study. In *Proc MONTECH Conf*, Montreal, Canada, 1987.
- 4 Zerhouni EA, Parish DM, Rogers WJ, Yang A and Shapiro EP: Human heart: tagging with MR imaging — a method for noninvasive assessment of myocardial motion. *Radiology*, 169 (1988) 59–63.
- 5 Hunter WC and Zerhouni EA: Imaging distinct points in left ventricular myocardium to study regional wall deformation. In *Innovation in Diagnostic Radiology* (Ed: JH Anderson), Springer, London, 1989, pp. 169–190.
- 6 Axel L and Dougherty L: MR Imaging of motion with spatial modulation of magnetization. *Radiology*, 171 (1989) 841–845.
- 7 Buchalter MB, Weiss JL, Rogers WJ, Zerhouni EA, Weisfeldt ML, Beyar R and Shapiro EP: Noninvasive quantification of left ventricular rotational deformation in normal humans using magnetic resonance imaging myocardial tagging. *Circulation*, 81 (1990) 1236–1244.
- 8 Boesiger PB, Fischer SE, Hess OM, Krayenbuhl HP, Maier SE and McKinnon G: Myocardial tagging for noninvasive quantification of myocardial motion. *Medica Mundi*, 36 (1991) 106–110.
- 9 Prinzen TT, Arts T, Prinzen FW and Reneman RS: Mapping of epicardial deformation using a video processing technique. *J Biomech.* 19 (1986) 263–273.

- 10 Prinzen FW, Augustijn CH, Arts T, Alessie MA and Reneman RS: Redistribution of myocardial fiber strain and blood by asynchronous activation, *Am J Physiol*, 259 (1990) H300–H308.
- 11 Augustijn CH, Arts T, Prinzen FW and Reneman RS: Mapping the sequence of contraction of the canine left ventricle, *Pflügers Arch*, 419 (1991) 529–533.
- 12 Prinzen FW, Augustijn CH, Alessie MA, Arts T, Delhaas T and Reneman RS: The time sequence of electrical and mechanical activation during spontaneous beating and ectopic stimulation, *Eur Heart J*, 13 (1992) 535–543.
- 13 Kong Y, Morris JJ and McIntosh HD: Assessment of regional myocardial performance from biplane coronary angiograms, *Am J Cardiol*, 27 (1971) 529–537.
- 14 MacKay SA, Potel ML and Rubin JM: Graphics methods for tracking three-dimensional heart wall motion, *Comput Biomed Res*, 15 (1982) 455–473.
- 15 Potel MJ, Rubin JM, MacKay SA, Aisen AM, Al-Sadir J and Sayre RE: Methods for evaluating cardiac wall motion in three dimensions using bifurcation points of the coronary arterial tree, *Invest Radiol*, 18 (1983) 47–57.
- 16 Young AA, Hunter PJ and Smaill BH: Estimation of epicardial strain using the motions of coronary bifurcations in biplane cineangiography, *IEEE Trans Biomed Eng*, 39 (1992) 526–531.
- 17 Ingels NB, Daughters GT, Stinson EB and Alderman EL: Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers, *Circulation*, 52 (1975) 859–867.
- 18 Ingels NB, Daughters GT, Stinson EB and Alderman EL: Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard, *Circulation*, 61 (1980) 966–972.
- 19 Garrison JB, Ebert WL, Jenkins RE, Yionoulis SM, Malcolm H, Heyler GA, Shoukas AA, Maughan WL and Sagawa K: Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms, *Comput Biomed Res*, 15 (1982) 76–96.
- 20 Hannaford B and Glantz SA: Adaptive linear predictor tracks implanted radiopaque markers, *IEEE Trans Biomed Eng*, 32 (1985) 117–125.
- 21 Ingels NB Jr, Daughters GT II, Stinson EB, Alderman EL and Miller DC: Three-dimensional left ventricular midwall dynamics in the transplanted human heart, *Circulation*, 81 (1990) 1837–1848.
- 22 Douglas AS, Rodriguez EK, O'Dell W and Hunter WC: Unique strain history during ejection in canine left ventricle, *Am J Physiol*, 260 (1991) H1596–H1611.
- 23 Niczyporuk MA and Miller DC: Automatic tracking and digitization of multiple radiopaque myocardial markers, *Comput Biomed Res*, 24 (1991) 129–142.
- 24 Rashid FR: Towards a system for the interpretation of moving light displays, *IEEE Trans Patt Anal Mach Intell*, 2 (1980) 574–581.
- 25 Yachida M, Asada M and Tsuji S: Automatic analysis of moving images, *IEEE Trans Patt Anal Mach Intell*, 3 (1981) 12–19.
- 26 Taylor KD, Mottier FM, Simmons DW, Cohen W, Pavlak R Jr, Cornell DP and Hankins GB: An automated motion measurement system for clinical gait analysis, *J Biomech*, 15 (1982) 505–516.
- 27 Ferrigno G and Gussoni M: Procedure to automatically classify markers in biomechanical analysis of whole-body movement in different sports activities, *Med Biol Eng Comput*, 26 (1988) 321–324.
- 28 Muijtjens AMM, Roos JMA, Prinzen TT, Hasman A, Reneman RS and Arts T: Noise reduction in estimating cardiac deformation from marker tracks, *Am J Physiol*, 258 (1990) H599–H605.
- 29 Stewart GW: *Introduction to Matrix Computations*, Academic Press, London, 1973, pp. 317–323.
- 30 Dixon WJ (Ed): *BMDP Statistical Software*, Univ of California Press, Berkeley, 1988, p. 813.
- 31 Gabriel KR and Zamir S: Lower rank approximation of matrices by least squares with any choice of weights, *Technometrics*, 21 (1979) 489–498.
- 32 *Numerical Algorithms Group FORTRAN Library Manual*, Mark 12, Vols. 4 and 5, Oxford (UK), 1987.
- 33 Little RJA and Rubin DB: *Statistical Analysis with Missing Data*, Wiley, New York, 1987.

Tracking markers with missing data by iterative lower rank approximation

Arno M.M. Muijtens, Jef M.A. Roos, Theo Arts*, Arie Hasman, and Robert S. Reneman**

Departments of Medical Informatics, Biophysics, and Physiology**,
Cardiovascular Research Institute Maastricht
University of Limburg, The Netherlands*

Submitted

Abstract — Motion and deformation of an object may be quantified by following attached markers in time, using 2 or more simultaneously recorded video or cine frame sequences. When using video (256×256 pixels, 50 s^{-1}) for recording cardiac motion, generally no more than approximately 20 markers can be followed simultaneously due to 1) missing of marker images causing tracks to be interrupted, and 2) close approach or crossing of marker tracks causing ambiguities in the identification of the marker images. In the present study we developed the Lower Rank Tracking (LRT) method which can automatically and reliably follow considerably more than 20 markers in a sequence of video frames. Initially conventional tracking is used to render subtracks with interruptions and unknown identification with markers. Then by using the Iterative Lower Rank (ILR) extrapolation method, subtracks are connected. In contrast with currently applied methods, the ILR method is intended to accurately extrapolate over wide gaps (up to 20 frames in a 40 frame sequence), thus avoiding ambiguities as much as possible. The performance of the LRT method was evaluated in computer simulations using tracks obtained by 2-D projection of markers attached to a naturally moving heart. To simulate realistic loss of data on marker image positions, generated marker images were skipped from further analysis when they mutually approached as close as 4 pixels. Using 75 markers in 60 frames, 90% of the markers was correctly followed, 10% could not be followed, and 0.3% of the tracks were false.

Introduction

In investigating cardiac mechanics, regional motion and deformation of the cardiac wall is measured in physiological experiments. A comparative evaluation [1] rendered that following radiopaque markers in 2 simultaneously recorded video or cine frame sequences in an X-ray set-up [2-6] is appropriate to quantify regional cardiac motion in physiological experiments.

When following markers or feature points, a major problem is the identification of the same marker in different frames, the so called correspondence problem [7,8]. With stereo observations, correspondences in time as well as between the two views must be established. The current study is restricted to correspondence in time. The track of a marker is given by the chain of measured positions in time (Fig. 1).

In reconstructing the tracks, often the similarity of intensity patterns is used to identify subimages which correspond to the same object point [9-11], but with identical markers only the positional information can be used for identification.

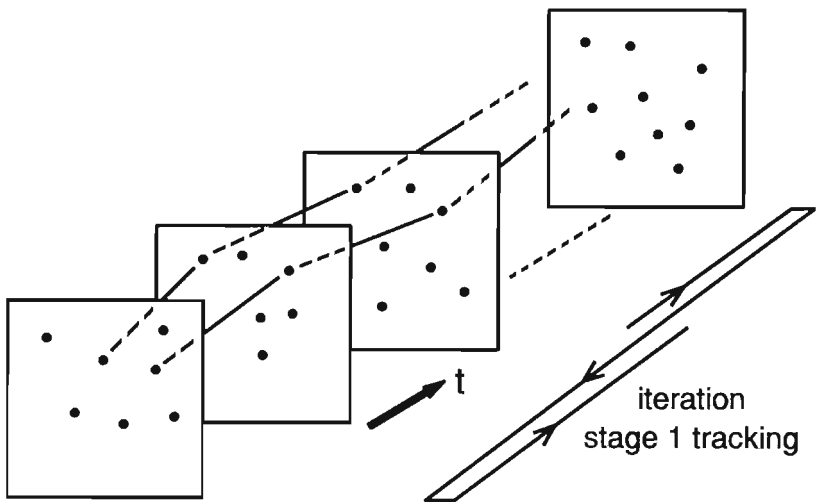


FIGURE 1. Schematic representation of marker image positions (dots) in a sequence of frames. The symbol t refers to time. A marker track is represented by a chain of marker image positions. The searching algorithm for chain formation starts somewhere in the middle of the sequence and moves back and forth through the sequence of frames.

Several constraints have been used to cope with the combinatorially explosive problem of finding corresponding marker images in a sequence of frames, such as maximum velocity, smoothness of motion, common motion, consistent match, and rigidity [12]. When the inter-frame marker displacements are small relative to the inter-marker distances, the corresponding marker image in the next frame is unambiguously found in a small region centred at the marker image position in the current frame. When the marker displacements increase or, equivalently, the inter-marker distances decrease, more frequently the search regions will overlap, resulting in ambiguity and false correspondences. Smaller regions and less failures may be achieved by centring the search region at an estimated extension of the track. When the motion is smooth, an estimate of the extension of a track in the next frame can be obtained by extrapolation of the last two or three positions [13-15]. However, this method is not appropriate for non-smooth displacements and extrapolation over time gaps of several frames. When observing cardiac motion at a video frequency of 50 s^{-1} , in certain episodes of the cycle smoothness of motion does not apply.

Occasionally, coherence of motion was used in predicting the extension of tracks. For instance in [15] the invariability of the distance between markers on rigid object parts was used and in [16] the similarity of motion of neighbouring markers. However, when radiopaque markers moving in 3-D are projected by X-ray into 2-D images, neighbouring marker images may have different motion patterns, because they correspond to markers which are distant in 3-D.

When using video (256×256 pixels, 50 s^{-1}) for recording cardiac motion, generally no more than approximately 20 markers can be followed simultaneously due to 1) missing marker images causing tracks to be interrupted, and 2) close approach or crossing of marker tracks causing ambiguities in the identification of the marker images [3,6,17]. The goal of the present study is to develop a method to reliably follow far more than 20 markers in a sequence of video frames. Then, markers must correctly be tracked through frequently occurring episodes with missing marker images or ambiguity about the identity of marker images.

Earlier [1], in following markers attached to the heart, we developed the Iterative Lower Rank (ILR) method which accurately extrapolates marker tracks over wide gaps in time. With 50 markers observed in 40 frames and up to 20 positions missing in a track, an extrapolation error of 0.5 pixel was obtained in tracks with a total marker excursion of approximately 20 pixels.

In the present study we have developed the Lower Rank Tracking (LRT) method to automatically reconstruct marker tracks by using the ILR extrapolation method in an iterative procedure. Initially parts of tracks (subtracks) are obtained by conventional tracking techniques, using two-point linear extrapolation [14] to estimate the extension of a track in an adjacent frame. Then corresponding subtracks are connected on the basis of applying ILR extrapolation to sets of subtracks.

The performance of the LRT method was evaluated in realistic computer simulations of markers attached to a moving left ventricle. Detection problems, which occur when

marker images partly overlap, were represented by deleting generated marker image positions when their mutual distance was less than a preset minimum distance. In the tracking procedure, the retained marker image positions became part of a correct or an incorrect track. A track is considered correct when it consists of marker image positions of only one physical marker. Besides being correct, the reconstructed tracks should also be of long duration, ideally, extending over the whole observation interval. The ability of the LRT method to connect large numbers of unidentified marker image positions to form correct long tracks was investigated for several numbers of markers, levels of the image position noise, and values of the minimum allowed inter-marker distance.

Tracking procedure

Object types

Table I shows the object types used in the tracking procedure, the corresponding symbols, indices and attributes. The motion of an object is quantified by following the movements of a set of M attached markers in a sequence of T cine or video frames. Marker detection in the t -th frame results in a set of unidentified marker image positions $P_{i,t}$, $i=1, \dots, I_t$ with the positions specified as image coordinates $(x,y)_{i,t}$. By conventional tracking, chains of marker image positions are identified that correspond to the same marker, resulting in a set of J subtracks, S_j , $j=1, \dots, J$. A subtrack is specified by t_b , t_e , the indices of the begin and end frame, respectively, and a list of indices i which refers to the associated marker image positions. In the second stage of the tracking procedure, ILR extrapolation is used to connect subtracks which correspond to the same physical marker. Thus a set of K reconstructed marker tracks is obtained. A reconstructed track is specified by a list of subtrack indices j . A marker image position may be associated to only one subtrack, and a subtrack may be part of only one reconstructed track. Furthermore, a subtrack contains a marker image position in each frame from time t_b until t_e , and subtracks which are part of the same reconstructed track may not overlap in time.

Partial tracking by conventional extrapolation

Starting at a frame in a period with minimum motion, a set of subtracks is constructed iteratively. The frame which has minimum m_t , with m_t the median inter-frame marker displacement defined in Eq. (9), Appendix A, is used as the starting frame. For each marker image position in the starting frame matching positions in the next frame are searched for. If two positions in the next frame match one position in the current frame, or vice versa, matching is considered ambiguous. Let d be the distance between a marker image position in frame 1 and a candidate match in the adjacent frame 2, then the match is only considered unambiguous if 1) d is less than a maximum distance r_1 , and 2) no other marker image position in frame 2 lies within a distance $r_2 \geq r_1$ from the image in frame 1. The criterion distances r_1 and r_2 are calculated for each frame according to their definitions presented in Appendix A.

TABLE I. Object types used in the tracking procedure.

Object type	Symbol	Index		Attribute
Marker image position ¹⁾	P	i, t	$i=1, \dots, I_t$ $t=1, \dots, T$	$(x, y)^{2)}$
Subtrack ³⁾	S	j	$j=1, \dots, J$	$t_b, t_e, \{i\}^{4)}$
Reconstructed track ⁵⁾	R	k	$k=1, \dots, K$	$\{j\}^{6)}$

¹⁾ The location of a marker image in frame t of a sequence of T frames

²⁾ Image coordinates of the marker image position

³⁾ Chain of marker image positions in consecutive frames, assumed to be associated with one physical marker

⁴⁾ Begin and end frame of the subtrack, and list of indices referring to the marker image positions in the subtrack

⁵⁾ Chain of subtracks, assumed to be associated with one physical marker

After completion of the search in a frame, tracking proceeds to the next frame. Then, subtracks obtained in former steps are linearly extrapolated as if marker velocity is constant, the search region is centred at the predicted marker position and the criterion distances r_1 and r_2 are calculated accordingly (see Appendix A). If for a marker image position matching is ambiguous or no match can be found, tracking is interrupted and the subtrack ends in the current frame. Unconnected marker image positions form the start of new subtracks. When the last frame has been reached, iteration is reversed in time towards the first frame, and then reversed again until the first analyzed frame has been reached (Fig. 1). So, in each frame tracking is applied in both time directions. If in one of the directions a match is ambiguous, the track is interrupted.

Connecting subtracks by ILR extrapolation

In an iterative procedure subtracks are connected to form reconstructed tracks. A time window is positioned at the beginning of the sequence of frames and the tracks which overlap the window are subdivided into two sets: long and short ones. By definition a long track has more than 50% detected marker image positions present in the current time window. Extrapolations are obtained by applying the ILR method to the set of long tracks. When a short track unambiguously matches the extrapolation of a long track, the two are connected to form a compound longer track. Several conditions should be satisfied when considering the connection of the long track R_{kl} with the short track $R_{k'2}$. The two tracks may not overlap in time and the distance between the extrapolations $Q_{kl,t}$

of the long track and the corresponding measured marker image position $P_{k_2,t}$ of the short track must be sufficiently small. The matching distance $\Delta(k_1, k_2)$ is defined

$$\Delta(k_1, k_2) = \overline{(Q_{k_1,t} - P_{k_2,t})^2} \quad (1)$$

where averaging is performed over available marker positions in the shorter track R_{k_2} and within the window. The match between tracks may be accepted if the matching distance defined in Eq. (1) satisfies

$$\Delta(k_1, k_2) < \delta_3 \quad (2)$$

with δ_3 a positive threshold value. The match of tracks R_{k_1} and R_{k_2} is considered unambiguous if there are no competing short tracks which overlap in time with R_{k_2} . If matches are competing, a match (R_{k_1}, R_{k_2}) is still considered unambiguous if the matching distance of the competing pair (R_{k_1}, R_{k_3}) is substantially larger than the distance of (R_{k_1}, R_{k_2}) , according to

$$\Delta(k_1, k_3) - \Delta(k_1, k_2) > \delta_4 \quad (3)$$

with δ_4 a positive constant.

Acceptance of a match (R_{k_1}, R_{k_2}) also depends on the number of observed marker positions f_C in the complete track R_{k_2} , the number of observed marker positions f_W of R_{k_2} within the window, and the number of missing observations f_G in the time gap between the tracks. Each match should also satisfy

$$\begin{aligned} & (f_W = 1 \text{ AND } f_C = 1) \text{ OR} \\ & (f_W = 2 \text{ AND } f_C \leq 3) \text{ OR} \\ & (f_W \geq 3 \text{ AND } (1 + f_G/f_W) \Delta(k_1, k_2) < \delta_4) \end{aligned} \quad (4)$$

Connecting subtracks starts with setting the window width to the initial value W_{init} and positioning the window at the beginning of the sequence of frames. The track connecting procedure is only applied within the current window when the number of long tracks is at least equal to $W_{init}/2$. Furthermore in each frame of the window, more than 50% of the long tracks should contain detected marker image positions. Repeatedly, the connecting procedure is carried out, followed by shifting the window over approximately 1/6 of its width. When the window reaches the last frame, the direction of

window shifting is reversed. Shifting back and forth is repeated until no additional connections can be made. Then the window is enlarged by one third, and the whole procedure is repeated. When finally the width of the window covers the entire observation interval, tracking is concluded.

Implementation

The tracking procedure was implemented in *FORTRAN* on a VAX8650 computer. In the evaluated implementation, the constants δ_1 and δ_2 (Eqs. (10) and (11), Appendix A) were set to 2.5 and 0.4, respectively. The threshold values δ_3 and δ_4 (Eqs. (2) and (3)) were set to 5 and 20 pixel², respectively, and the initial window width W_{init} was set to 8. The run-time of the procedure ranged from 3 to 15 minutes for 60 frames of computer simulated data of 25 to 100 markers.

Performance evaluation

Computer simulations of markers attached to a moving left ventricle

Marker image positions were generated by using a model for 3-D motion of the left ventricle. In this model the motion of the wall of the left ventricle is described by 13 time-varying parameters [18]. In the referred study estimates of the kinematic parameters were obtained for a biplane cine recording of 14 radiopaque myocardial markers in 60 frames sampled at 90 s⁻¹ in an animal experiment. The resulting parameter values were used as a basis for the computer simulations in the present study. In the reference state the ventricular wall was defined between two truncated confocal ellipsoids (Fig. 2), in accordance with the model used in [19]. In the 3-D motion field, as defined by the 13 time-varying kinematic parameters, the ventricle was oriented with the long axis coinciding with the torsion axis. The angle between the camera axis and the ventricular long axis was about 75°. In agreement with a laboratory set-up the image distance between the camera source point and the image plane was chosen to be 1200 mm, and the object distance was about 600 mm. Markers were positioned at random in the ventricular wall with the restriction that the minimum inter-marker distance d_{min} must be complied with:

$$d_{min} = 0.9 \times 2r$$

$$\frac{4}{3} \pi r^3 = \frac{V_{wall}}{M} \quad (5)$$

with M the number of markers to be positioned and V_{wall} the volume of the ventricular wall. If a marker position did not comply with the minimum distance condition, a new

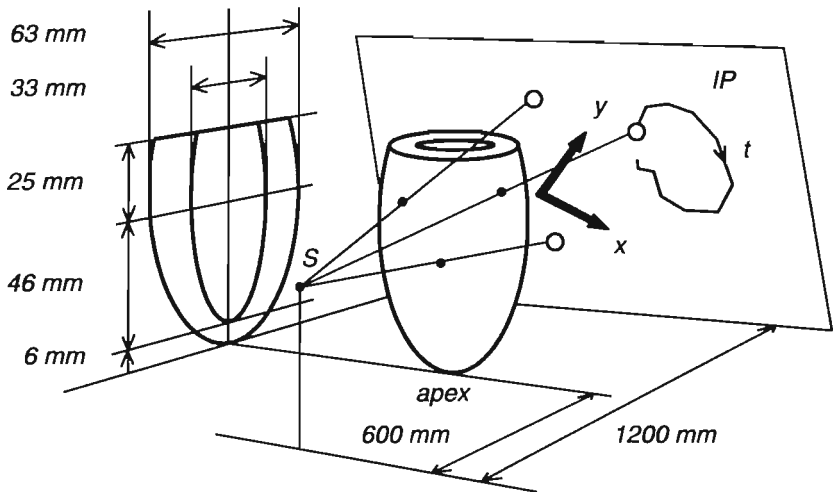


FIGURE 2. Computer simulation of an X-ray camera set-up. The left ventricle of the heart is positioned half-way between the X-ray focus S and the image plane IP . The simulated ventricular wall is rotationally symmetric and defined between two truncated confocal ellipsoids, as indicated at the left. The ventricle moves according to a kinematic model with motion parameters varying in time as measured in a physiological experiment. Radiopaque markers are randomly distributed across the ventricular wall. The projection of a marker in the plane IP follows a path in time, which is sampled at 90 s^{-1} . The symbols x , y and t refer to image coordinates and time, respectively.

random position was tried repeatedly until the condition was satisfied. Throughout the cardiac cycle marker positions were obtained at a frame rate of 90 s^{-1} using parameter values as determined experimentally on the basis of the kinematic model [18]. The corresponding 2-D marker image positions were obtained by the perspective projection from the source point on the image plane, as indicated in Fig. 2. Finally, uncorrelated gaussian noise with a root mean square (RMS) value σ was added to the generated marker image position coordinates. When circular optical markers of 1.5 mm diameter attached to the ventricular wall were observed in frames of 256×256 pixels, σ was found to be approximately 0.1 pixel [20]. In the current study the standard RMS noise level was taken 5 times as high, in order to account for the loss of accuracy due to the lower contrast and motion blurring in a set-up with continuous X-ray.

Detection problems which occur when the X-ray shadow images of markers overlap were simulated by deleting marker image positions if the mutual distance d was below a given minimum distance d_m :

$$d \leq d_m \quad (6)$$

Evaluation of the LRT method

Computer simulated data were generated for the number of markers M set equal to 25, 50, 75, 100, and the RMS noise level σ set equal to 0.25, 0.5 and 1.0 pixel. For each combination of M and σ the results were obtained in 10 repeated simulations. In each simulation a new random configuration of markers and a new random noise pattern were generated. Missing data were introduced according to condition (6), with the minimum allowed marker image distance d_m set to 2, 4, and 8 pixels, respectively.

The tracking procedure has been evaluated by calculating several indices, expressing the quality of the reconstruction. Starting with M markers in T frames (Fig. 3), in the simulation the number of marker images is diminished to n_R due to close approach of marker images (condition (6)). A reconstructed track is only correct if all associated marker image positions correspond to the same physical marker. Furthermore, given a threshold D_m in duration of a track, the set of reconstructed tracks has been subdivided into short and long tracks. The number of marker image positions belonging to correct long tracks (n_{CL}) has been calculated as a function of the duration threshold D_m . Similarly, the number of marker image positions in false long tracks (n_{FL}) has been calculated. From these calculations two indices of performance were derived. Given a threshold value $D_m = 50$ frames for $T = 60$ frames and referring to the number of retained markers n_R , the index f_{CL} expresses the fraction of marker images belonging to correct long tracks:

$$f_{CL} = n_{CL}(50) / n_R \quad (7)$$

with $n_{CL}(50)$ representing the value of n_{CL} for $D_m=50$ frames. The other index M_{CL} expresses the average number of marker image positions per frame which is part of a correct long marker track:

$$M_{CL} = n_{CL}(50) / T \quad (8)$$

Ideally, the values of f_{CL} and M_{CL} should be equal to 1 and M , respectively. When increasing the number of markers, a maximum is to be expected in the value of M_{CL} , because of the increasing number of missing marker images and the reduced proportion of long tracks rendered by the tracking procedure.

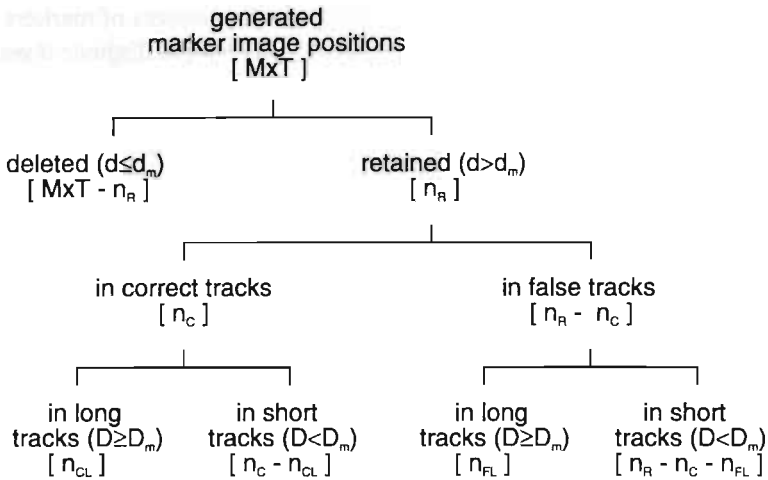


FIGURE 3. Handling of generated marker image positions in a computer simulation for tracking. The number of elements per category is indicated between square brackets, starting with $M \times T$, with M and T the number of markers and frames, respectively. Positions are considered missing when their mutual distance d is smaller than a given distance d_m . A reconstructed track is considered correct when all positions in the track correspond to the same physical marker. Reconstructed tracks are subdivided into the categories of long and short tracks by the threshold D_m of the duration D of a track.

Results

An example of a set of generated marker tracks is presented in Fig. 4 for a computer simulation with 75 markers in 60 frames. Fig. 5 shows the fraction of marker image positions in correct (f_{cor}) and false (f_{fal}) long tracks as a function of the duration threshold D_m . Results with an RMS noise level σ of 0.25 pixel and a minimum allowed distance between marker image positions d_m of 2 pixels are not shown because they were similar to the results obtained with $\sigma=0.5$ pixel and $d_m=4$ pixels, respectively. In panels a,b and e,f, the faster descending curves correspond to the tracks obtained by conventional tracking in the first stage of the procedure. Most of the tracks are broken into subtracks thus limiting the duration. The slowly descending curves correspond to the reconstructed tracks obtained at the end of the tracking procedure. Then subtracks have been connected by ILR extrapolation, and as a result the fraction of tracks with longer duration is substantially increased. Different line types indicate the numbers of markers used in the simulations. In all simulations only in the second stage of the tracking procedure rarely a false connection between subtracks was established (panels c,d,g,h).

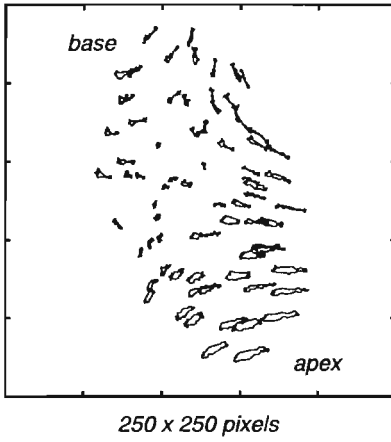


FIGURE 4. Computer generated tracks of radiopaque markers observed in an X-ray camera set-up. Tracks of 75 markers attached to a simulated moving left ventricle follow loops during the cardiac cycle.

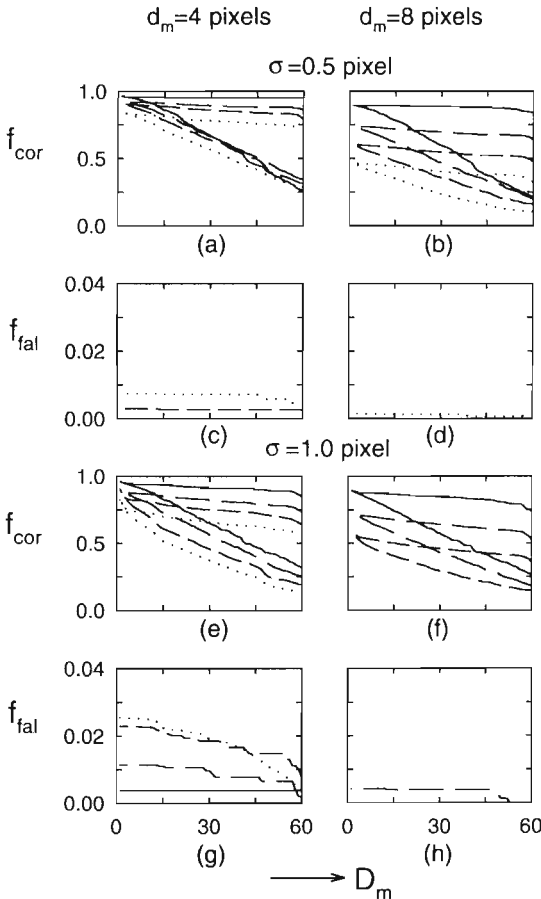


FIGURE 5. Results of the Lower Rank Tracking method in computer simulations of markers attached to a moving heart. The fraction of generated marker image positions in the correct (f_{cor}) and false tracks (f_{fal}) are plotted as a function of the track duration threshold D_m . The results of the initial conventional tracking and of the complete tracking procedure are presented in panels a,b,e,f (fastly and slowly descending curves, respectively). Results are shown for simulations with an RMS noise level σ of 0.5 (panels a,b,c,d) and 1.0 pixel (panels e,f,g,h), with values of the minimum distance allowed between marker image positions, d_m , of 4 (left panel) and 8 pixels (right panel), and a number of markers equal to 25, 50, 75, and 100 (solid, long dash, medium dash, dots).

TABLE II. Results of the LRT method in computer simulations of markers attached to a moving left ventricle.

$\sigma^{1)}$ pixel	$d_m^{2)}$ pixel	$M=25^{3)}$		$M=50$		$M=75$		$M=100$	
		$f_{DE}^{4)}$	$f_{CL}^{5)}$	f_{DE}	f_{CL}	f_{DE}	f_{CL}	f_{DE}	f_{CL}
0.5	4	4	99	7	95	9	91	14	88
0.5	8	10	97	25	90	38	84	52	79
1.0	4	3	92	7	84	9	78	15	71
1.0	8	10	90	25	80	40	69	-	-

¹⁾ RMS noise level of the marker image position measurement

²⁾ Minimum distance allowed between marker image positions

³⁾ Number of markers used in the simulations

⁴⁾ The percentage of the $M \times 60$ in 60 frames generated marker image positions, which was deleted because positions were mutually closer than d_m

⁵⁾ The percentage of the retained marker image positions which belonged to correct reconstructed tracks of a duration of at least 50 frames

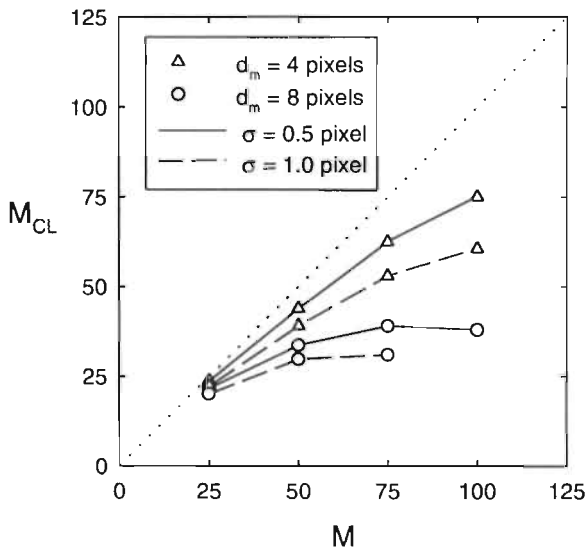


FIGURE 6. The performance index M_{CL} is the average number of marker image positions per frame belonging to a correct track with a duration of at least 50 frames. M_{CL} is shown as a function of the number of markers M . The computer simulations were carried out with a minimum distance allowed between marker image positions, d_m , of 4 (triangle) and 8 pixels (circle), and the RMS noise level σ equal to 0.5 (solid line) and 1 pixel (dashed line), respectively. The dotted line $M_{CL}=M$ refers to perfect performance.

Table II shows f_{DE} , the percentage of generated marker image positions which was deleted because of a mutual distance smaller than d_m , and f_{CL} , the percentage retained positions in correct long tracks (≥ 50 frames) as defined in Eq. (7). The average number per frame of positions belonging to correct long tracks, M_{CL} defined in Eq. (8), is shown in Fig. 6 as a function of the number of markers. The dotted line indicates the theoretical optimum value of the index which is obtained when all marker image positions are detected and are correctly followed in tracks of a duration of 60 frames.

Fig. 7 shows in more detail the results of the tracking procedure for a simulation with 75 markers, a noise level of 0.5 pixel and a minimum distance of 4 pixels allowed between marker images. Each row represents the 60 marker image positions of a complete marker track. Some tracks were mutually close in long episodes, and as a result have many positions missing (blanks). Subtracks (light horizontal lines) obtained by conventional tracking were correctly connected (bold horizontal lines) by the use of ILR extrapolation. Rarely a false connection between subtracks of different physical markers was made (diagonal bold line).

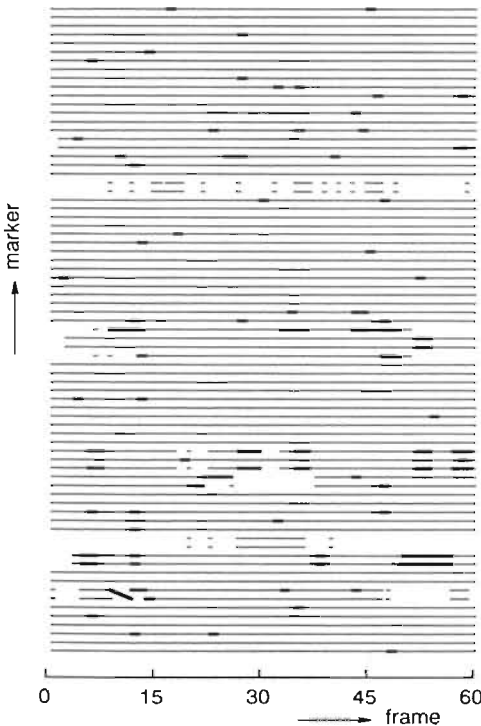


FIGURE 7. Schematic representation of the tracking results obtained with the LRT method in a computer simulation of 75 left ventricular markers observed in a sequence of 60 frames. The noise level was 0.5 pixel and the minimum distance allowed between marker image positions was equal to 4 pixels. Each row represents the 60 marker image positions of a complete marker track. Observed and missing marker image positions are indicated by short horizontal lines and blanks, respectively. The light horizontal lines indicate subtracks which were established by the initial conventional tracking, and bold lines indicate the connections obtained by ILR extrapolation in the second stage of the tracking procedure. The diagonal bold line represents a false relation between two subtracks.

An example of two crossing marker tracks is shown in Fig. 8. The two marker tracks (upper panel, circles and triangles) approached very closely in frames 14 to 16, resulting in missing marker image positions (squares in upper panel). By applying the LRT method correct subtracks (middle panel, solid lines) and connections (dotted lines) were obtained. ILR extrapolation applied to the final set of reconstructed tracks resulted in estimates for the missing positions (lower panel, circles and triangles) which closely resemble the true tracks (solid lines). The example shows that, despite the complication of crossing marker tracks, correctly reconstructed tracks and accurate estimates of missing marker image positions can be obtained by the iterative use of ILR extrapolation as implemented in the LRT method.

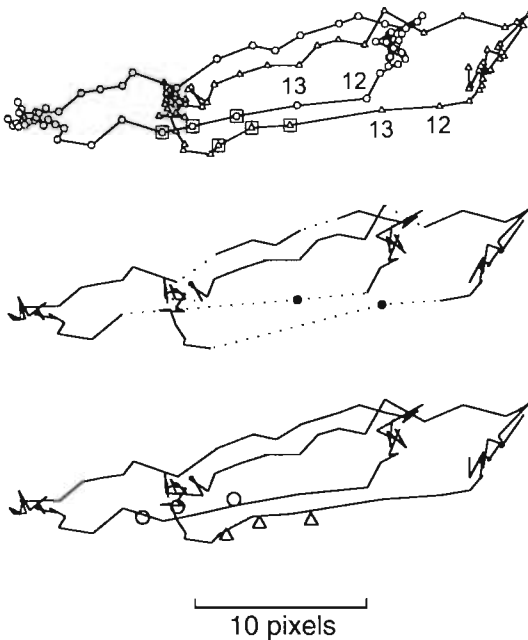


FIGURE 8. Example of crossing marker tracks in a computer simulation with 50 markers, a noise level of 0.5 pixel, and a minimum distance of 4 pixels allowed between marker image positions. The two marker tracks are indicated in the upper panel (frames 12 and 13 indicated) by circles and triangles. The tracks approach too close in frames 14 to 16 resulting in missing marker image positions (open squares). Tracking of the unconnected marker images by the LRT method results in subtracks (middle panel, solid lines) which by the use of ILR extrapolation are correctly connected (dotted lines). Missing marker image positions can be estimated quite accurately (lower panel, circles and triangles) by applying ILR extrapolation to the final set of reconstructed tracks. For comparison the complete true tracks are also shown (solid lines).

Discussion

When quantifying motion and deformation of an object by following attached identical markers in a sequence of video frames (256×256 pixels, 50 s^{-1}), a major problem is the identification of corresponding marker images in consecutive frames. With conventional extrapolation methods applied in tracking markers attached to heart, generally no more than approximately 20 markers can be followed simultaneously [3,6,17]. With more markers major complications are the interruptions of marker tracks due to detection failures, or due to the ambiguity which arises when tracks approach very closely or cross in the projection.

In the present study we developed the Lower Rank Tracking (LRT) method to reliably follow considerably more than 20 markers with part of the marker images occasionally being missing. In the first stage of the procedure, by conventional tracking, subtracks are obtained with interruptions and unknown identification. Decisions about ambiguous situations are postponed until the second stage, when more reliable decisions can be obtained by using the distance between two sequences of marker image positions, rather than the distance between two single marker image positions. At this stage, the ILR extrapolation method [1] is used to connect subtracks which are associated with the same physical marker.

The performance of the procedure was evaluated in computer simulations of markers attached to a moving left ventricle. Detection problems were represented by deleting marker image positions when their mutual distance was less than a preset minimum distance. With constants δ_1 and δ_2 (Eqs. (10) and (11), Appendix A) set to 2.5 and 0.4, respectively, conventional tracking resulted in subtracks of reasonable length (Fig. 5, panels a,b,e,f, faster descending curves), while no false subtracks were established. With the threshold values δ_3 and δ_4 (Eqs. (2) and (3)) set to 5 and 20 pixel², respectively, only rarely a false match was established between subtracks (Fig. 5, panels c,d,g,h), and 70% to 99% of the retained marker image positions became part of correct long tracks (Table II).

The criterion for a match between two tracks is a small distance between extrapolations and correspondingly observed marker image positions, as defined by Eq. (1). The criterion is more reliable when the number of frames of the overlap is larger. By condition Eq. (4) the decision about a connection is postponed when the part of the overlap within the window is still narrow. Also, consecutive tracks which are separated by a narrow time gap are favoured by Eq. (4) in order to reduce the probability of establishing a false connection.

When investigating the motion of markers attached to the heart, ILR extrapolations of reasonable accuracy can be obtained with a set of at least 4 tracks, and with per frame and per track less than 50% marker image positions missing [1]. To fulfill these requirements with the many short tracks resulting from the conventional tracking in the first stage of the procedure (Fig. 5, panels a,b,e,f), subtracks are connected iteratively, starting in a window of minimum width $W_{\text{init}}=8$, which is iteratively shifted and enlarged while the tracks grow.

In evaluating the performance of the tracking procedure two indices were investigated: the fraction of detected marker image positions belonging to correct long tracks and the average number of marker image positions per frame in correct long tracks, f_{CL} and M_{CL} defined in Eqs. (7) and (8), respectively. For a noise level of 0.5 pixel and the minimum allowed distance between marker images, d_m , set to 4 pixels (Table II, first row), the procedure performed well: with the number of markers varying from 25 to 100, f_{CL} ranged from 99% to 88% while the fraction deleted positions (f_{DE}) ranged from 4% to 14%. Doubling the noise level resulted in a decrease of f_{CL} to values ranging from 92% to 71% (Table II, row 3). When also d_m was doubled to 8 pixels, the fraction deleted positions considerably increased resulting in reduced tracking performance (Table II, row 4). Still, with 75 markers 69% of the retained positions was finally part of a correct long track, despite the 40% deleted positions. However, with these extreme conditions, a number of 100 markers was beyond the limits of performance of the procedure.

For increasing numbers of markers, performance index M_{CL} deteriorated increasingly (Fig. 6, optimum M_{CL} indicated by dotted line). With a minimum allowed distance d_m of 8 pixels the index even started to decrease when the number of markers was beyond 75. With $d_m=4$ pixels, however, even with more than 75 markers the index still ascends, implying that an increase of markers still results in an increase of appropriate motion data. Substitution of missing marker image positions, e.g. by ILR approximation [1] might even further improve these results.

In a stereo camera set-up, a stereo pair of marker images is lost if either of the two marker images is lost. Hence, the expected fraction of appropriate pairs is equal to the squared fraction of appropriate marker images per projection. As a consequence, M_{CL} as a function of M will level off sooner for stereo pairs. For 100 markers, $d_m=4$ pixels and $\sigma=0.5$ pixel the value of M_{CL} is 56 and for $\sigma=1.0$ pixel and 75 markers a maximum M_{CL} value of 38 is reached. The calculations indicate that for numbers of markers beyond 75 the number of appropriate stereo pairs hardly increases or may even decrease.

Conclusions

The tracks of large numbers of markers can automatically and reliably be reconstructed by using the LRT method. In computer simulations 75 markers attached to a moving left ventricle were observed in 60 frames (256×256 pixel images, RMS position noise of 0.5 pixel). Marker images were considered missing, when they mutually approached closer than 4 pixels. Thus 10% of the generated positions was deleted. From the retained positions, a fraction f_{CL} of 90% became part of correct long tracks (duration ≥ 50 frames). A fraction of 10% could not be followed and 0.3% was part of a false track.

Performance index M_{CL} was defined as the average number of marker image positions per frame which is part of a correct long track. Ideally it should be equal to the number

of markers M , but it was found to level off for increasing M . For $M > 75$, M_{CL} increased moderately ($d_m = 4$ pixels), or even decreased ($d_m = 8$ pixels), indicating that motion information cannot be increased anymore by increasing the number of markers.

When increasing the RMS noise level σ or the minimum allowed inter-marker distance d_m , both performance measures, f_{CL} and M_{CL} , decrease. With $25 \leq M \leq 100$, doubling of σ from 0.5 to 1.0 pixel results in a 6% to 20% decrease of both f_{CL} and M_{CL} . Doubling of d_m from 4 to 8 pixels results in a decrease of f_{CL} by 2% to 10%, but in a much greater loss of M_{CL} by 10 to 50% due to the increased amounts of missing positions. A reduction of σ from 0.5 to 0.25 pixel, or of d_m from 4 to 2 pixels did not have a significant effect on performance.

Appendix A

Matching criteria in constructing subtracks

When calculating the criterion distances r_1 and r_2 , two types of marker images are distinguished, those included in a subtrack and single ones. For the two types, the search region in the next frame is centred at a position obtained by extrapolation of the subtrack, or at the marker position in the current frame, respectively. The size of the search region is based on an estimate of the inter-frame marker displacement. Let $h_{i,t}$ be the distance from the i -th marker image position in frame t to the nearest marker image position in the next frame. Then the median inter-frame displacement in frame t is defined

$$m_t = \text{median} \{ h_{i,t}, i = 1, \dots, L_t \} \quad (9)$$

with L_t the number of single marker image positions in frame t . Let m'_t be similarly defined as the median inter-frame displacement between predicted and observed marker image positions. Then, for an arbitrary marker image position P in the current frame t the distance r_1 is defined

$$r_1 = \begin{cases} \delta_1 m'_t, & \text{if } P \text{ in a track} \\ \delta_1 m_t, & \text{else} \end{cases} \quad (10)$$

with δ_1 a positive constant. Let the search region for a match with P in the next frame be centred at point Q , and let P' be the candidate match for P , then distance r_2 is defined

$$r_2 = \max (r_1, |QP'| \cdot \delta_2 r_1) \quad (11)$$

with δ_2 a positive constant.

References

- [1] A.M.M. Muijtjens, J.M.A. Roos, T. Arts, A. Hasman, and R.S. Reneman, "Extrapolation of incomplete marker tracks by lower rank approximation," *Int. J. Biomed. Comput.* **33**, 219-239 (1993).
- [2] J.B. Garrison, W.L. Ebert, R.E. Jenkins, S.M. Yionoulis, H. Malcom, G.A. Heyler, A.A. Shoukas, W.L. Maughan, and K. Sagawa, "Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms," *Comput. Biomed. Res.* **15**, 76-96 (1982).
- [3] W.C. Hunter and E.A. Zerhouni, "Imaging distinct points in left ventricular myocardium to study regional wall deformation," In: J.H. Anderson, *Innovations in Diagnostic Radiology*, pp. 169-190 (Springer, London, 1989).
- [4] N.B. Ingels, G.T. Daughters, E.B. Stinson, and E.L. Alderman, "Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers," *Circulation*, **52**, 859-867 (1975).
- [5] N.B. Ingels, G.T. Daughters, E.B. Stinson, and E.L. Alderman, "Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard," *Circulation*, **61**, 966-972 (1980).
- [6] M.A. Niczyporuk and D.C. Miller, "Automatic tracking and digitization of multiple radiopaque myocardial markers," *Comput. Biomed. Res.* **24**, 129-142 (1991).
- [7] Y. Yakimovsky and R. Cunningham, "A system for extracting three-dimensional measurements from a stereo pair of TV cameras," *Comp. Graph. Im. Proc.* **7**, 195-210 (1978).
- [8] J.K. Aggarwal and N. Nandhakumar, "On the computation from sequences of images - A review," *Proc. IEEE*, **76**, 917-935 (1988).
- [9] S.T. Barnard and W.B. Thompson, "Disparity analysis of images," *IEEE Trans. Patt. Anal. Mach. Intell.* **2**, 333-340 (1980).
- [10] M. Yachida, M. Asada, and S. Tsuji, "Automatic analysis of moving images," *IEEE Trans. Patt. Anal. Mach. Intell.* **3**, 12-19 (1981).
- [11] D.S. Kalivas and A.A. Sawchuk, "A region matching motion estimation algorithm," *CVGIP Image Understanding*, **54**, 275-288 (1991).
- [12] D.H. Ballard and C.M. Brown, *Computer Vision*, (Prentice-Hall, Englewoods Cliffs, NJ, 1982).
- [13] F.R. Rashid, "Towards a system for the interpretation of moving light displays," *IEEE Trans. Patt. Anal. Mach. Intell.* **2**, 574-581 (1980).
- [14] K.D. Taylor, F.M. Mottier, D.W. Simmons, W. Cohen, R. Pavlak Jr., D.P. Cornell, and G.B. Hankins, "An automated motion measurement system for clinical gait analysis," *J. Biomech.* **15**, 505-516 (1982).
- [15] G. Ferrigno and M. Gussoni, "Procedure to automatically classify markers in biomechanical analysis of whole-body movement in different sports activities," *Med. Biol. Eng. Comput.* **26**, 321-324 (1988).
- [16] T.T. Prinzen, T. Arts, F.W. Prinzen, and R.S. Reneman, "Mapping of epicardial deformation using a video processing technique," *J. Biomech.* **19**, 263-273 (1986).
- [17] A. Macerata, W.J. Sanders, E.L. Alderman, and D.C. Miller, "An automated video image analysis for detection and tracking implanted myocardial markers," In: *Proc. IEEE Conf. Comput. Cardiol. Washington 23-25 Sept. 1988*, pp. 547-550 (IEEE Computer Society Press, Washington, 1989).

- [18] T. Arts, W.C Hunter, A. Douglas, A.M.M. Muijtjens, and R.S. Reneman, "Description of the deformation of the left ventricle by a kinematic model," *J. Biomech.* **25**, 1119-1127 (1992) .
- [19] P.H.M. Bovendeerd, T. Arts, J.M. Huyghe, D.H. van Campen, and R.S. Reneman, "Dependence of local left ventricular wall mechanics on myocardial fiber orientation: a model study," *J. Biomech.* **25**, 1129-1140 (1992).
- [20] A.M.M. Muijtjens, J.M.A. Roos, T.T. Prinzen, A. Hasman, R.S. Reneman, and T. Arts, "Noise reduction in estimating cardiac deformation from marker tracks," *Am. J. Physiol.* **258**, H599-H605 (1990).

3-D Reconstruction of marker positions from stereo images using ML estimation

Arno M.M. Muijtjens, Jef M.A. Roos, Theo Arts*, and Arie Hasman

Departments of Medical Informatics, and Biophysics,
Cardiovascular Research Institute Maastricht
University of Limburg, The Netherlands*

Submitted

Abstract —Motion and deformation of the cardiac wall may be measured by following the motion of implanted radiopaque markers in 3-D, using two X-ray cameras simultaneously. Regularly, calibration of the position measurement system is obtained by registration of the images of a calibration object, containing 10-20 radiopaque markers at known positions. Unfortunately, an accidental change of the position of a camera after calibration requires complete recalibration. In the present study a method is presented by which calibration is obtained from the images during measurement, thus avoiding a separate calibration procedure. A camera is represented by a source point and an axis through this point, directed perpendicularly to the image plane. The stereo camera set-up is described by 5 dimensionless parameters and a scaling distance defining the relative position and orientation of the cameras. The parameters were estimated by a maximum likelihood (ML) method, and by conventional nonlinear least squares (NLS) methods. In a computer simulation n markers were attached to the surface of a spherical object, which was observed in a stereo camera set-up with perpendicular camera axes, and a source point distance $d_s = 10^3$ mm. The object was imaged on video (256×256 pixels), and for both cameras 200 pixels corresponded to the object aperture a of 0.1 rad. The position noise was set to an RMS (root mean square) value, σ , of $\pm 2.5 \times 10^{-4}$ rad (0.5 pixel). With the ML method the parameters were estimated without noticeable bias. For a sphere of about 70 mm diameter the positions of the markers were reconstructed ($n=25$) with an accuracy of ± 1 mm (RMS). In general, for markers attached to the surface of a sphere, the reconstruction error was found to be equal to $[2 d_s (n-5)^{-1/2} \sigma a^{-1}]$. When using multiple frame sequences in time, the accuracy of the calibration increased because the number n increased proportionally. Conventional NLS estimation showed considerable bias in one of the parameter estimates, and was less accurate than the ML method. Reconstructions based on the ML method were also more accurate than the results reported for linear least squares estimation (LLS) in an 8 parameter linear model.

Introduction

Regional motion and deformation of the cardiac wall may be assessed in physiological experiments by following implanted radiopaque markers [1-5] or bifurcation points of the coronary arterial tree [6-9]. The 3-D positions of the markers were reconstructed from stereo pairs of projections, obtained by simultaneously operated X-ray cameras. Comparable set-ups were described for measuring deformation of the lung [10-12], the diaphragm [13], and the pharynx [14-16]. A similar analysis was applied to the projections of optical markers to measure human motion [17], or deformation of biological tissues [18,19].

Reconstruction of the position of a marker from a stereo pair of projections requires knowledge of the geometry of the camera set-up. In an orthogonal camera set-up the position of the cameras can be determined directly, but slight inaccuracies in the related measurements reduce the accuracy of the 3-D reconstruction considerably [1,6]. The reconstruction is facilitated by the use of a calibration object, having a set of markers at precisely known positions. Before or after the measurement the object is placed in the measuring volume. Using the projections of the calibration object, the geometric parameters of the stereo camera set-up can be calculated [4,5,7,20,21]. This information is then used to reconstruct the 3-D position of a marker. The most prominent disadvantage of the method is that after calibration the cameras should maintain their positions exactly. Else a complete recalibration is needed, meaning removal of the object to be measured from its position between the cameras to enable repositioning of the calibration object.

The position of a marker is defined by 3 coordinates. A stereo projection of a marker provides 4 coordinates. So, each marker provides redundant information, which commonly is used to pair marker images [3,21].

When using a calibration object, the position and orientation of two cameras are calculated with respect to the coordinate system as attached to the calibration object. The viewing projection of a 3-D point to the image plane is described as a linear transformation, using a 3×4 matrix of 11 unknown elements and one element set equal to 1. Per view, a marker of the calibration object provides 2 constraining equations, hence, a minimum number of 6 markers is required to solve for the unknown 11 matrix elements. Generally, the number of markers in the calibration object is taken larger (12-20) in order to improve the accuracy by statistical averaging [3-5,7].

In absence of a calibration object, the coordinate system should be attached to the system of cameras. The relative position and orientation of the cameras is defined by 6 parameters, 3 translational and 3 rotational coordinates [22]. A camera can only measure angles, and no distances. So, apart from an unknown scaling distance, only 5 parameters are needed to define the relative orientation of the cameras. The redundant information conveyed by the fourth coordinate of the stereo projection of a marker may be used to estimate the geometry of the stereo camera set-up. So, at least 5 markers have to be imaged for the estimation of the 5 set-up parameters.

Several methods are presented to estimate the geometry of the stereo camera set-up.

Generally, the number of parameters used to describe the geometry is larger than 5, resulting in inherent redundancy in the parameter set used. Eight parameters were used in a linear estimation method [23], which was evaluated and further improved [24-26]. Less redundant nonlinear models were also introduced [27,28], and improvement of 3-D position measurement was obtained by using a multidimensional nonlinear optimization algorithm [29].

In the present study, a nonlinear model of the geometry of a stereo camera set-up was based on the absolute minimum of 5 parameters. Without the use of a calibration object these parameters were estimated from stereo pairs of images of a set of markers. A new maximum likelihood (ML) method was developed in which bias in the estimates was avoided, and the accuracy of the position reconstruction was improved. Because of its intended use in measuring deformation of the heart, the model is developed for an aperture angle of the object of 0.1 rad.

In computer simulations markers were distributed randomly over the surface of a sphere. The accuracy and bias of the estimated geometric parameters and the reconstructed marker positions were assessed. For comparison a more conventional nonlinear least squares (NLS) estimation technique was evaluated. In reconstructing the position of the markers, the performance of an earlier described linear least squares (LLS) technique based on 8 parameters [25] was also evaluated.

Geometric model of a stereo camera set-up

Assuming an ideal pinhole model [30], a camera is represented by a source point S through which all imaging rays pass, and an axis through S , directed perpendicularly to the image plane, intersecting the plane in C (Fig. 1). An image P_i of object point P is represented by 2-D coordinates (u,v) . The camera axis SC is defined to be of unit length, so image coordinate u is identical with the tangent of the angle between the camera axis SC and the line from S to the image plane position $(u,0)$, and a similar interpretation holds for coordinate v . Knowledge of the image P_i restricts the location of the corresponding object point P to the ray SP_i . When an additional image of P is obtained by a second camera, the location of P can be restricted to a single point.

The upper panel of Fig. 2 shows a stereo camera set-up with a left (L) and right (R) camera, the respective source points S_L and S_R , unit-length camera axes $S_L C_L$ and $S_R C_R$, and 2-D image coordinate systems (u_L, v_L) , and (u_R, v_R) . Since the scale is arbitrary, without loss of generality, the baseline distance $S_L S_R$ may be set to one. Points S_L' and S_R' , are the intersections of the baseline $S_L S_R$ with the right and left image planes, respectively.

A rectangular 3-D coordinate system, attached to the stereo camera system, is defined for reference purposes. Define camera axis angles

$$\omega_L = \angle C_L S_L S_R \quad \omega_R = \pi - \angle C_R S_R S_L \quad (1)$$

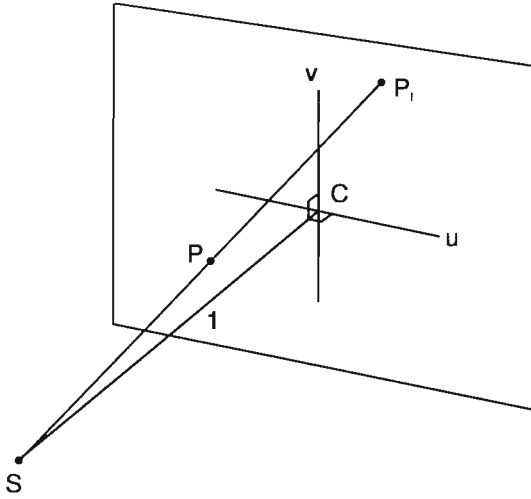


FIGURE 1. Pinhole camera model represented by source point S , and unit length camera axis SC , which is perpendicular to the image plane. Image coordinate system u, v centred in C . Image P_i and the corresponding object point P .

and let O be the point in the bisector of the planes $S_L S_R C_L$ and $S_L S_R C_R$, which satisfies

$$\angle OS_L S_R = \omega_L \quad \angle OS_R S_L = \pi - \omega_R \quad (2)$$

Then define the xyz -coordinate system with the origin in O , the x -, and y -axis in the plane $S_L S_R O$, parallel with and perpendicular to the baseline $S_L S_R$, respectively (Fig. 2, upper panel).

Apart from a scaling distance, the geometry of the stereo camera set-up is defined by 5 angular parameters,

$$\begin{aligned} \varphi_L &= \angle(\mathbf{e}_{uL}, \text{plane } S_L S_R C_L) \\ \varphi_R &= \angle(\mathbf{e}_{uR}, \text{plane } S_L S_R C_R) \\ \nu &= (\omega_R - \omega_L)/2 \\ \tau &= (\omega_R + \omega_L - \pi)/2 \\ \gamma &= \angle(\text{plane } S_L S_R C_L, \text{plane } S_L S_R C_R) \end{aligned} \quad (3)$$

where \mathbf{e}_{uL} and \mathbf{e}_{uR} are the unit vectors in the direction of the indicated coordinates. The angles ω_L and ω_R are replaced by angles ν and τ for the mathematical convenience of symmetry. The positive direction of angle γ is taken to the right around $\overline{S_L S_R}$, the vector pointing from S_L to S_R .

The reconstruction of object point P from a corresponding stereo pair (P_L, P_R) , is defined by specifying the xyz-coordinates $(\hat{x}, \hat{y}, \hat{z})$ of point \hat{P} the estimation of P , as a function of the observed image coordinates $\{u_L, v_L, u_R, v_R\}$ and the parameters of the set-up. Let angles α_L, β_L , and α_R, β_R be defined

$$\begin{aligned}\alpha_L &= \angle(S_R S_L P_L) & \beta_L &= \angle(\text{plane } S_L S_R C_L, \text{plane } S_L S_R P_L) \\ \alpha_R &= \pi - \angle(S_L S_R P_R) & \beta_R &= \angle(\text{plane } S_R S_L C_R, \text{plane } S_R S_L P_R)\end{aligned}\quad (4)$$

with the positive direction of the elevation angles β_L and β_R taken to the right around vector $\overline{S_L S_R}$. In general, due to the noise the image rays $S_L P_L$ and $S_R P_R$ do not exactly intersect, and point P is estimated at a point in the vicinity of both rays. In our approach, point P is reconstructed in the bisector of the left and right imaging ray planes, $S_L S_R P_L$ and $S_R S_L P_R$, at the point \hat{P} which satisfies

$$\begin{aligned}\angle(\hat{P} S_L S_R) &= \alpha_L \\ \angle(\hat{P} S_R S_L) &= \pi - \alpha_R \\ \angle(\text{plane } z=0, \text{plane } S_L S_R \hat{P}) &= (\beta_L + \beta_R)/2\end{aligned}\quad (5)$$

The angles α_L, β_L , and α_R, β_R in Eq. (4) are defined as functions of the image coordinates $\{u_L, v_L, u_R, v_R\}$ and the set-up parameters by the relations

$$\begin{aligned}|S_L P_L|^2 &= 1 + (u_L^*)^2 + (v_L^*)^2 & |S_R P_R|^2 &= 1 + (u_R^*)^2 + (v_R^*)^2 \\ |P'_L P''_L| &= \sin \omega_L - u_L^* \cos \omega_L & |P'_R P''_R| &= \sin \omega_R - u_R^* \cos \omega_R\end{aligned}\quad (6)$$

$$\begin{bmatrix} u_L^* \\ v_L^* \end{bmatrix} = \begin{bmatrix} \cos \varphi_L & \sin \varphi_L \\ -\sin \varphi_L & \cos \varphi_L \end{bmatrix} \begin{bmatrix} u_L \\ v_L \end{bmatrix}; \quad \begin{bmatrix} u_R^* \\ v_R^* \end{bmatrix} = \begin{bmatrix} \cos \varphi_R & \sin \varphi_R \\ -\sin \varphi_R & \cos \varphi_R \end{bmatrix} \begin{bmatrix} u_R \\ v_R \end{bmatrix}\quad (7)$$

where (u_L^*, v_L^*) and (u_R^*, v_R^*) are the coordinates of images P_L and P_R with respect to the rotated image coordinate systems (Fig. 2, middle panel). Distance $P'_L P''_L$ is defined in the same figure, and $P'_R P''_R$ is similarly defined for the right camera. Angles ω_L and ω_R are defined in Eq. (3) as functions of parameters ν and τ . The specification of the reconstruction is completed by defining the coordinates $(\hat{x}, \hat{y}, \hat{z})$ as a function of the angles α_L, β_L , and α_R, β_R ,

$$\begin{aligned}
\hat{x} &= |S_L \hat{P}''| - |S_L O'| & \hat{y} &= |\hat{P} \hat{P}''| \cos \hat{\beta} - |OO'| & \hat{z} &= |\hat{P} \hat{P}''| \sin \hat{\beta} \\
|S_L \hat{P}''| &= |\hat{P} \hat{P}''| \operatorname{ctg} \alpha_L & |S_L O'| &= |OO'| \operatorname{ctg} \omega_L & (8) \\
|\hat{P} \hat{P}''| &= (\operatorname{ctg} \alpha_L - \operatorname{ctg} \alpha_R)^{-1} & |OO'| &= (\operatorname{ctg} \omega_L - \operatorname{ctg} \omega_R)^{-1} \\
\hat{\beta} &= (\beta_L + \beta_R)/2
\end{aligned}$$

where the indicated distances are defined in the lower panel of Fig. 2. In order to actually perform the reconstruction, the parameters of the stereo camera set-up must be known.

Estimation of the parameters of the stereo camera set-up

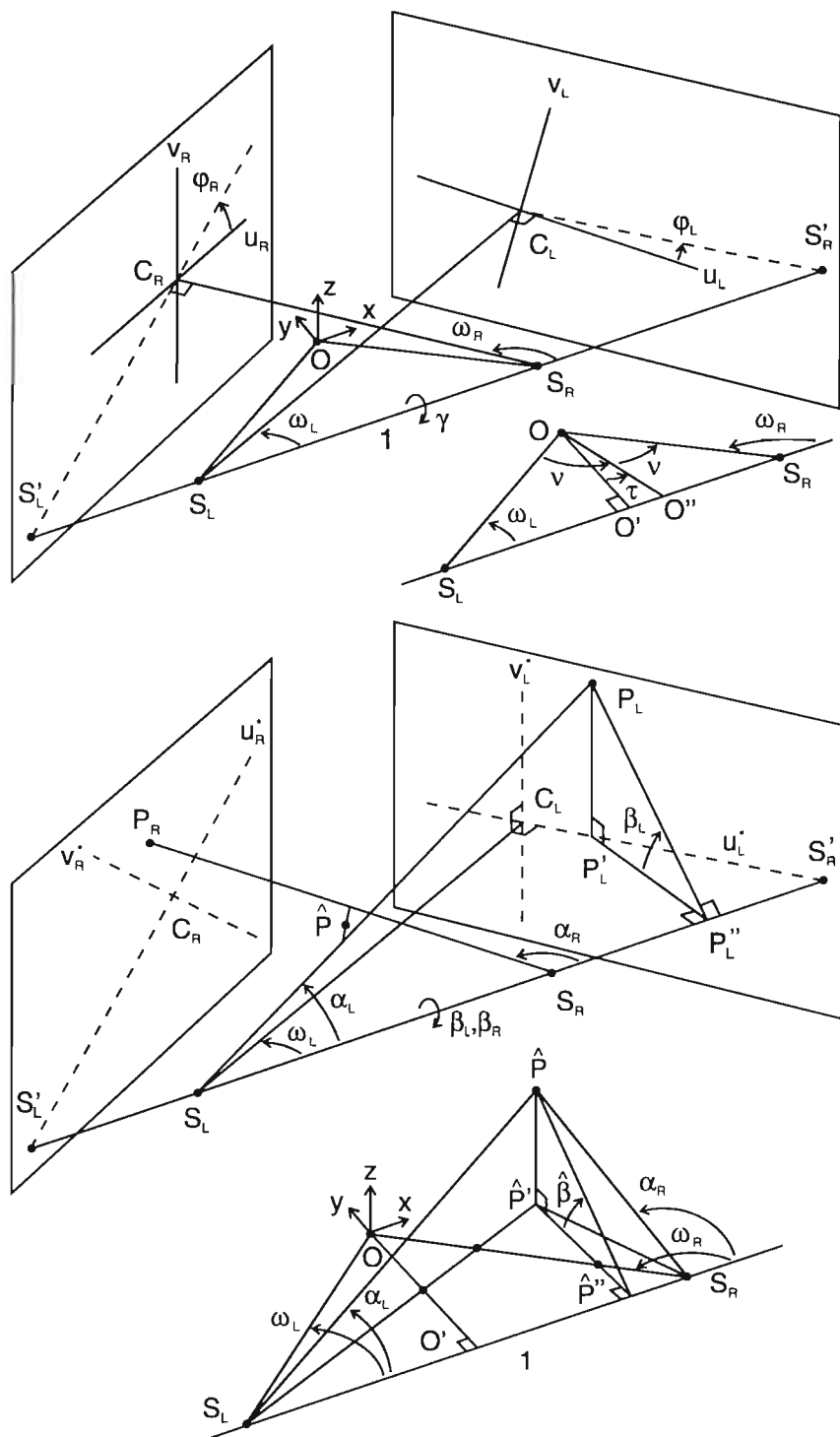
In the model defined above, the elevation angle difference between the left and right imaging ray planes,

$$\Delta \beta = \beta_L - \beta_R - \gamma \quad (9)$$

is used as a measure of stereo mismatch. When n stereo pairs are available, the 5×1 vector of set-up parameters

$$\theta = (\varphi_L, \varphi_R, v, \tau, \gamma)^T \quad (10)$$

FIGURE 2. Geometric model of a stereo camera set-up. *Upper panel:* Ideal pinhole cameras. Source points S_L and S_R , at distance unity. Unit length camera axes $S_L C_L$ and $S_R C_R$, perpendicular to the image planes with coordinate systems $u_L v_L$, $u_R v_R$, respectively. Points S_L' and S_L'' intersections of baseline $S_L S_R$ with the image planes. Set-up parameters, angles φ_L , φ_R , v , τ , and γ . Coordinate system with origin O in bisector of planes $S_L S_R C_L$ and $S_L S_R C_R$, and x and y axes parallel with and perpendicular to $S_L S_R$, respectively. *Inset:* Relation between angles ω_L, ω_R , and v, τ . *Middle panel:* Points P_L and P_R , the noise perturbed images of the unknown object point P . Point \hat{P} , the reconstruction of P in the bisector of planes $S_L S_R P_L$ and $S_L S_R P_R$. Angle α_L , the angle of $S_L P_L$ and $S_L S_R$. Elevation angle β_L , the angle between the plane $S_L S_R P_L$ and the plane $S_L S_R C_L$. Angles α_R and β_R similarly defined for the right camera. Points P_L' and P_L'' , the projections of P_L on the u_L' -axis, and on the baseline, respectively. *Lower panel:* Coordinates of point \hat{P} as function of angles $\alpha_L, \beta_L, \alpha_R, \beta_R, \omega_L, \omega_R$. Point \hat{P}' the projection of point \hat{P} on the plane $z=0$. Points O' and \hat{P}'' , the projection on the baseline of points O , and \hat{P} , respectively. Angle $\hat{\beta}$, the elevation angle of point \hat{P} with respect to the plane $z=0$.



can be estimated by minimizing the sum of squares of the elevation angle differences. The corresponding estimate is expressed

$$\hat{\theta}_{ang} = \theta \mid \sum_{i=1}^n \Delta \beta_i^2 \text{ minimum} \quad (11)$$

with $\Delta \beta_i$ the mismatch for the i -th stereo pair, and the right-hand part meaning: the value of vector θ which minimizes the indicated function.

An alternative estimate may be found by minimizing the differences in the elevation arc length,

$$\hat{\theta}_{arc} = \theta \mid \sum_{i=1}^n \Delta \beta_i^2 b_i^2 \text{ minimum} \quad (12)$$

with b_i equal to the distance $|\hat{P}\hat{P}''|$ defined in the lower panel of Fig. 2, for the i -th stereo pair. Hence, in Eq. (12) a weighted sum of squares of elevation angle differences is minimized.

In Appendix A an error analysis is reported, resulting in the specification of weights, w_i^2 , which are inversely proportional to the variance of the error in $\Delta \beta_i$. When gaussian errors are assumed, the corresponding weighted least squares estimate is the maximum likelihood (ML) estimate,

$$\hat{\theta}_{ML} = \theta \mid \sum_{i=1}^n \Delta \beta_i^2 w_i^2 \text{ minimum} \quad w_i^2 = (r_{L,i}^{-2} + r_{R,i}^{-2})^{-1} \quad (13)$$

with $r_{L,i}$, the distance $P_L'P_L''$ indicated in the middle panel of Fig. 2, for the i -th stereo pair, and $r_{R,i}$ similarly defined for the image P_R in the right image plane.

The estimates defined in Eqs. (11)-(13), which are obtained by applying standard nonlinear least squares (NLS) procedures (routine F04GDF of the NAG Fortran Library [31]), will be referred to by *NLS angle*-, *NLS arc*-, and *ML method*. With $n=100$ stereo pairs, the parameter estimation took less than 1 s run-time on a VAX8650.

3-D Reconstruction error

The position of a 3-D object point, P_i , $i=1, \dots, n$, and its reconstruction, \hat{P}_i , is specified by position vectors p_i and \hat{p}_i , respectively, and the corresponding displacement error vector is defined

$$\mathbf{e}_{dpl,i} = \hat{\mathbf{p}}_i - \mathbf{p}_i \quad (14)$$

The reconstruction of n points attached to an object, results in a displaced and deformed version of the original object, and can be described by the linear displacement model

$$\hat{\mathbf{p}}_i = \mathbf{h} + F\mathbf{p}_i + \delta_i, \quad i=1,\dots,n \quad (15)$$

where \mathbf{h} is a 3×1 translation vector, F is 3×3 linear deformation matrix and δ_i is a 3×1 vector of the residual displacement which is not explained by the linear model. The matrix F can be decomposed (Appendix B) according to

$$F = sRG \quad (16)$$

with scalar s a scale factor, R a 3×3 orthogonal rotation matrix, and G a 3×3 matrix representing a deformation of an incompressible object.

Given a set of n pairs $(\hat{\mathbf{p}}_i, \mathbf{p}_i)$, $i=1,\dots,n$, the elements of vector \mathbf{h} and matrix F were estimated by least squares on the basis of Eq. (15), and the scalar s , and matrices R and G were calculated (Appendix B). Removal of the displacements corresponding to uniform translation, rotation, and change of scale, according to

$$\mathbf{e}_{def,i} = s^{-1} R^T (\hat{\mathbf{p}}_i - \mathbf{h}) - \mathbf{p}_i \quad (17)$$

resulted in the error vector $\mathbf{e}_{def,i}$ which accounts for the pure deformation error of the reconstruction. By substitution of Eqs. (15) and (16) into Eq. (17), the deformation error was further decomposed

$$\begin{aligned} \mathbf{e}_{def,i} &= \mathbf{e}_{ldf,i} + \mathbf{e}_{res,i} \\ \mathbf{e}_{ldf,i} &= G \mathbf{p}_i - \mathbf{p}_i \quad \mathbf{e}_{res,i} = s^{-1} R^T \delta_i \end{aligned} \quad (18)$$

with error vectors $\mathbf{e}_{ldf,i}$ and $\mathbf{e}_{res,i}$ representing the uniform linear deformation error, and the residual deformation error, respectively. The magnitude of the errors defined in Eqs. (14),(17) and (18) is quantified by the root mean square (RMS) level of the vector lengths. The corresponding RMS values are indicated ε_{dpl} , ε_{def} , ε_{ldf} , and ε_{res} , respectively. If not otherwise indicated, in the current study the term ‘reconstruction error’ refers to the deformation error defined in Eq. (17).

The random reconstruction error, e_{ran} , represents the error due to only the image position noise. It corresponds to the displacements found for exact set-up parameters and noisy image positions. The random error represents the inevitable, minimum value of the reconstruction error which is not affected by the accuracy of the parameter estimation.

Performance evaluation

The performance of the parameter estimation and the 3-D reconstruction was evaluated by mathematical analysis and in computer simulation experiments. Three situations were evaluated:

- A. Markers attached to the surface of a sphere,
- B. Markers attached to a simulated moving left ventricle (LV),
- C. Markers distributed over the volume of a sphere.

The major part of the evaluation was based on *Situation A*.

For the standard X-ray stereo camera set-up used in the evaluation it holds

$$(\varphi_L, \varphi_R, v, \tau, \gamma) = (0, 0, \frac{\pi}{4}, 0, 0) \quad (19)$$

so, camera axes were perpendicular and intersecting, and object distances were equal. The object was imaged on video (256×256 pixel). For both cameras 200 pixels corresponded to the object aperture u of 0.1 rad, u being defined as the ratio of the object diameter, and the object distance to the source point. With the distance between the two camera source points, d_s , set to 10^3 mm, the diameter of the observed object was approximately 70 mm. The image position noise was set to an RMS value, σ , of 2.5×10^{-4} rad (0.5 pixel), a value which was motivated previously [32]. The number of markers M was set to 25. For the single frame observations in *Situations A and C*, the number of stereo pairs n was equal to M . In *Situation B* the heart motion was observed in a sequence of T frames, so n was equal to $M \times T$.

Situation A: Markers attached to the surface of a sphere

In the standard stereo camera set-up described above, markers were distributed at random across the surface of a sphere with its centre in point O (Fig. 2). By mathematical analysis and in computer simulations, two aspects were investigated

1. Parameter bias, and parameter and reconstruction accuracy,
2. Reconstruction error induced per parameter.

1. Parameter bias, and parameter and reconstruction accuracy

Mathematical analysis of the parameter estimation

For a set of object points which is homogeneously distributed across the surface of a sphere, in Appendix C it was derived that the estimation error in the parameters τ , ϕ_L and ϕ_R , is approximately proportional to a^{-1} , the reciprocal of the aperture. The error is proportional to a^{-2} for v , and the estimate of γ is unaffected by variations of a . Also, the estimates of the 5 parameters were found to be uncorrelated.

Computer simulations

Generation of data. A number of $M=25$ markers was positioned at random on the spherical surface, with the restriction that a minimum inter-marker distance must be complied with. The minimum distance was defined

$$d_{\min} = 1.4 R (\pi/M)^{1/2} \quad (20)$$

with R the radius of the spherical object. The value of d_{\min} defined in Eq. (20) is based on the most dense and the least dense homogeneous distribution of markers given d_{\min} . A polygon with sides equal to d_{\min} is associated with both distributions. For the most dense one the polygon consists of triangles and a marker is located at each corner, and for the least dense one the polygon consists of hexagons with a marker located in the centre. The corresponding occupied area per marker is equal to d_{\min}^2 and $3 \times d_{\min}^2$, respectively, and Eq. (20) is obtained by multiplying the compromise value $2 \times d_{\min}^2$ with the number of markers M and setting the result equal to the available area $4\pi R^2$.

Projection of the 3-D marker positions resulted in exact image coordinates, and after adding uncorrelated gaussian noise with an RMS level σ , a set of $n=M$ observed stereo image pairs $\{u_L, v_L, u_R, v_R\}_i$, $i=1, \dots, n$ was obtained. Data were generated for several combinations of values of the variables n , σ , and a . A change of aperture was obtained by changing the scale of the object.

Calculation of bias and RMS error. The parameter bias, which is equal to the average parameter estimation error, and the RMS value of the parameter estimation error and the reconstruction error were obtained in 100 repeated simulations. In each repeated simulation, the random configuration of markers and the random noise were newly generated. Initial estimates for the parameters were obtained by random sampling in intervals centred at the true values of the parameters defined in Eq. (19), with interval width $\pm 10^\circ$ for parameters ϕ_L , ϕ_R , v , and τ , and $\pm 2^\circ$ for γ . The chosen interval widths represent the prior uncertainty in the 5 stereo camera set-up parameters. After estimation of the parameters, the corresponding reconstruction error was accurately obtained in a separate simulation with 100 markers distributed across the spherical surface. The systematic displacement error induced by the parameter errors was calculated by using the parameter estimates and the exact image positions. The resulting errors were

TABLE I. Number of stereo pairs n , RMS position noise σ , and aperture a in the computer simulations used in the comparison of conventional NLS and the ML method.

computer simulations	n	σ 10^{-4} rad	a rad
a	25	5.0	0.1
b	25	2.5	0.1
c	100	2.5	0.1
d	25	2.5	0.2

decomposed according to Eqs. (14)-(18). The random reconstruction error defined in the Section '3-D Reconstruction error' was calculated by using the exact parameter values and the noisy image positions.

ML method and conventional NLS methods. The three estimation procedures, the NLS angle-, NLS arc- and ML method defined in Eqs. (11)-(13), were applied to identical sets of simulated data, and the parameter bias and the RMS parameter and reconstruction error were obtained. Data were generated for the values of n , σ , and a shown in Table I. As a reference, the value of statistical significance in a two sided t-test of bias=0 at a p -level of 5%, was calculated.

Accuracy of ML method as a function of n , σ and a . The RMS parameter and reconstruction errors were obtained in simulations with n , σ , and a set to the standard values, and also for n equal to 10, and 100, and for σ , and a set to twice, and half the standard values of 2.5×10^{-4} , and 0.1 rad, respectively.

2. Reconstruction error induced per parameter

Mathematical analysis. The systematic reconstruction error induced by the parameter errors, was derived for each parameter separately, and was decomposed according to Eqs. (14)-(18). The linear deformation part of the reconstruction error, represented by matrix G in Eq. (16), was shown (Appendix D) to correspond to lengthening and shortening of the object in two orthogonal directions, proportional to the magnitude of the parameter error (Fig. 3). For a homogeneously distributed set of markers on the surface of a sphere with radius R , it was derived (Appendix D) that errors Δv in parameter v , $\Delta \tau$ in τ , and $\Delta \phi$ in ϕ_L or ϕ_R result in the respective RMS linear deformation errors

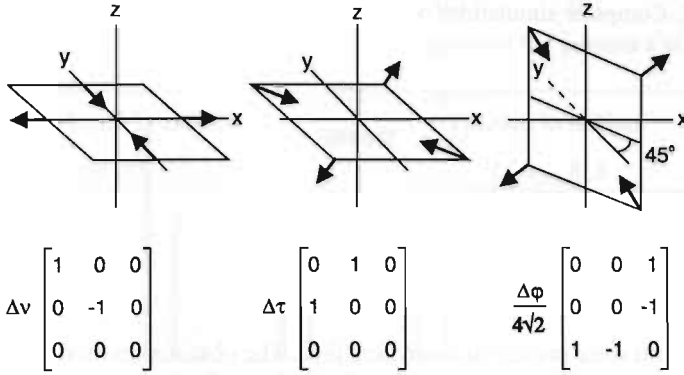


FIGURE 3. Deformation error in 3-D object reconstruction. The parameters were separately perturbed. For markers attached to surface of a sphere the mathematically derived deformation is indicated in the xyz -coordinate system. Also the corresponding matrix $G-I$ is shown, with 3×3 matrices G and I , the incompressible deformation matrix, and the unit matrix, respectively. From left to right: reconstruction error induced by errors Δv , $\Delta \tau$, and $\Delta \phi$, in parameters v , τ , and ϕ_L , respectively: lengthening and shortening along the x - and y -axis, along axes $x = \pm y$, $z = 0$, and in directions in the $x=y$ plane, at angles of 45° and 135° with the horizontal, respectively.

$$\varepsilon_{ldf,v} = c R |\Delta v| \quad \varepsilon_{ldf,\tau} = c R |\Delta \tau| \quad \varepsilon_{ldf,\phi} = c R |\Delta \phi/4| \quad c = (2/3)^{1/2} \quad (21)$$

The results for parameters ϕ_L or ϕ_R are identical because of the symmetry of the set-up, and therefore are indicated with ϕ . An error in parameter γ does not induce any erroneous displacements.

Computer simulations. With one of the set-up parameters set to a perturbed value, the error displacements of 100 markers were calculated, and decomposed according to Eqs. (14)-(18). The calculations were performed for aperture values of 0.05, 0.1 and 0.2 rad. The parameter perturbations were set to the RMS parameter error found with the ML method in the simulations in *Situation A*, for 25 markers, the standard noise level, and similar aperture values.

Situation B: Markers attached to a simulated moving left ventricle

Computer generated data. A simulated left ventricle (LV) was positioned at point O with the base-apex axis coinciding with the z -coordinate (Fig. 2, upper panel). The ventricular geometry, the random positioning of M markers in the LV wall, and the

TABLE II. Computer simulations of multiple frame stereo observations of markers attached to a moving left ventricle.

number of markers	number of frames
5, 8, 10, 25	1, 10, 25, 50

simulated motion were presented elsewhere [32]. The obtained motion data agreed with real LV motion, measured in a dog at a frame rate of 90 s^{-1} in a sequence of $T=50$ frames, with the aperture and the noise level set to the standard values.

Performance evaluation ML method. The RMS parameter errors were calculated for the combinations of numbers of markers and frames shown in Table II. For $T < 50$, the available 50 frames were sampled with a frequency $T/50$, e.g. with $T=10$, the frames at sample times 5, 10, 15, ..., and 50 in the sequence were used.

Situation C: Markers distributed over the volume of a sphere

Data were generated according to the specifications in Fencil and Metz [25]. In the standard set-up defined by Eq. (19), markers were randomly scattered over the volume of a sphere, which was observed at an object aperture of 0.18 rad . For $n=8, 10$, and 30 , uniformly distributed pixel round off noise of an amplitude of $1.8 \times 10^{-4} \text{ rad}$ was added. Also, for $n=10$, gaussian noise of RMS levels $\sigma=3.5 \times 10^{-4}$, and $7.0 \times 10^{-4} \text{ rad}$ (1 and 2 pixels in a 512×512 image) was added, supplemented by the pixel round off error. The ML method was applied to the generated data, and the RMS deformation error was calculated according to Eqs. (14)-(18).

Results

Situation A: Markers attached to the surface of a sphere

ML method and conventional NLS methods. The resulting bias and parameter accuracy of parameter v varied substantially for the three estimation methods defined in Eqs. (11)-(13). For the other parameters no significant bias was found and the parameter accuracy of the three methods was similar. The upper panel of Fig. 4 shows the absolute

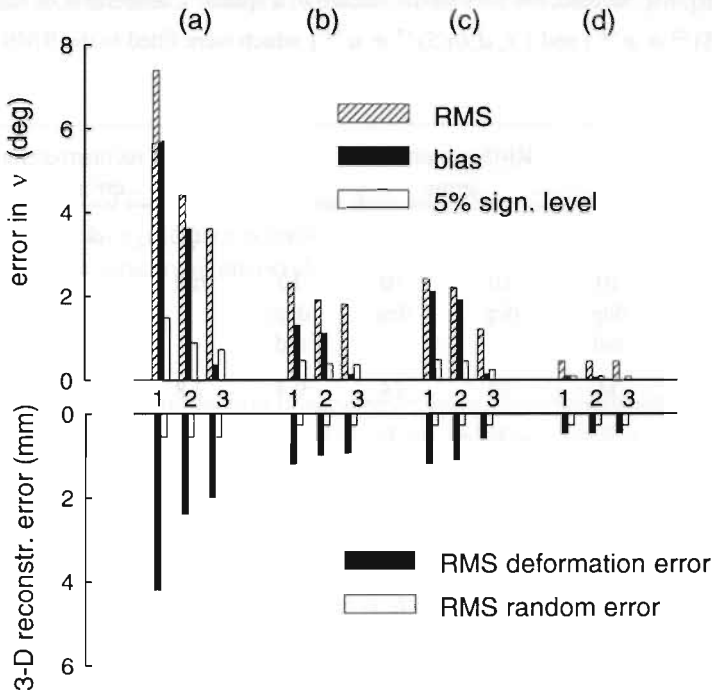


FIGURE 4. Estimation bias and accuracy for angle v , and accuracy of the 3-D reconstruction. Computer simulations of markers attached to the surface of a sphere. Estimation methods 1) NLS angle, 2) NLS arc, and 3) ML, correspond to minimization of the sum of squared elevation angle differences, elevation arc differences, and optimally weighted elevation arc differences, respectively. Data were simulated for the number of stereo pairs n , the RMS image position noise σ , and the aperture a , respectively, set to 25, 5×10^{-4} , 0.1 (panel a), 25, 2.5×10^{-4} , 0.1, (panel b), 100, 2.5×10^{-4} , 0.1 (panel c) and 25, 2.5×10^{-4} , 0.2 (panel d). Results were based on 100 repeated simulations per condition. *Upper panel*: Parameter bias, RMS error, and 5% significance level of bias=0. *Lower panel*: 3-D reconstruction deformation error due to the parameter errors, and the random error due to the image position noise.

value of the bias in parameter v (solid bar) for the simulation conditions (indicated a, b, c, d) presented in Table I. For comparison, the RMS parameter error (dashed bar) and the 5% significance level for the bias (blank bar) are also shown. The highly significant value of the bias for the NLS angle- and NLS arc method (1,2 of column a) persisted when the noise level σ was reduced from 5.0×10^{-4} to 2.5×10^{-4} rad (column b), or the number of stereo pairs n was increased from 25 to 100 (column c). Only when the aperture a was increased, from 0.1 to 0.2 rad, the bias decreased to a nonsignificant level (column d). For the ML method (3 in columns a to d) the estimate of v was

TABLE III. ML method, markers on the surface of a sphere. Coefficients of functions $[f_K(n-5)^{-1/2} \sigma a^{-q_K}]$ and $[f_K d_s(n-5)^{-1/2} \sigma a^{-q_K}]$ which were fitted to the RMS errors.

K	RMS parameter error				RMS reconstruction error		
	ν	φ	τ	γ	dpl	def	res
unit of f_K	10 deg rad	10 deg	10 deg	10 deg rad	rad		rad ⁻¹
f_K	34	28	14	9.4	7.8	2.0	1.4
q_K	2	1	1	0	2	1	0

unbiased, more accurate, and, as a consequence, the reconstruction error ϵ_{def} was minimum (lower panel of Fig. 4).

Accuracy of the ML method as a function of n , σ and a . In the upper panel of Fig. 5, the RMS parameter estimation errors, ϵ_ν , ϵ_φ , ϵ_τ , and ϵ_γ , are presented as a function of n , σ , and a . The lower panel of Fig. 5 shows the RMS values of the resulting reconstruction errors, ϵ_{dpl} , ϵ_{def} , ϵ_{res} , and ϵ_{ran} defined in Eqs. (14)-(18) and the accompanying text. The RMS parameter error in a nonlinear least squares fit is expected [35] to be proportional to σ , the RMS value of the noise in the observations, and inversely proportional to the square root of $n-5$, the number of degrees of freedom of the residual sum of squares. Accordingly, in order to summarize the results, curves were LS fitted to the RMS error values in Fig. 5, using the functions

$$\begin{aligned}
 f_K (n-5)^{-1/2} \sigma a^{-q_K} & \quad K \in \{\nu, \varphi, \tau, \gamma\} \\
 f_K d_s (n-5)^{-1/2} \sigma a^{-q_K} & \quad K \in \{dpl, def, res\}
 \end{aligned}
 \tag{22}$$

where label K indicates a parameter or one of the reconstruction errors defined in Eqs. (14), (17), and (18), f_K is a constant, and exponent q_K was set to a value corresponding to Eq. (42), Appendix C, or a value found by inspection of Fig. 5. The source point distance d_s was introduced as a scale factor. In Table III the resulting values for q_K and

f_k are shown. For the RMS random reconstruction error, which solely depends on σ , it holds

$$\varepsilon_{ran} = 1.1 d_s \sigma \quad (23)$$

In the log-log plot of Fig. 5, the functions defined in Eqs. (22) and (23) are represented by straight lines. In agreement with the results in Appendix C, the estimates of the 5 parameters were found to be uncorrelated.

Reconstruction error as a function of parameter error. For the standard stereo camera set-up defined in Eq. (19), the radius R of the spherical object can be written $d_s \cos(\pi/4) a/2$. When substituted in Eq. (21), for the RMS linear deformation errors induced per parameter it holds

$$\begin{aligned} \varepsilon_{ldf,v} &= c_v d_s a |\Delta v| & \varepsilon_{ldf,\tau} &= c_\tau d_s a |\Delta \tau| & \varepsilon_{ldf,\phi} &= c_\phi d_s a |\Delta \phi| \\ c_v = c_\tau &= 5.0 \times 10^{-3} \text{ rad}^{-1} \text{ deg}^{-1} & c_\phi &= 1.25 \times 10^{-3} \text{ rad}^{-1} \text{ deg}^{-1} \end{aligned} \quad (24)$$

with a in radian, and the parameter errors Δv , $\Delta \tau$, and $\Delta \phi$ in degrees. Application of the decomposition defined in Eqs. (14)-(18) to the error displacements in the simulations, resulted in matrices which agreed with Appendix D.

Fig. 6 (upper panel) shows the ratio of the RMS deformation error defined in Eq. (17), and the parameter error which caused it, as a function of the aperture. The RMS linear deformation error (Fig. 6, solid lines), calculated according to Eq. (24), was found to be almost identical to the RMS deformation error (difference $\pm 2\%$). The lower panel of Fig. 6 shows the ratio of the RMS residual reconstruction error defined in Eq. (18), and the RMS deformation error, as a function of the aperture.

ML method, parameter errors and deformation error for the standard situation. In Table IV the major results for the standard values $n=25$ and $\sigma=2.5 \times 10^{-4}$ are summarized. The contribution per parameter to the RMS linear deformation error was calculated by applying Eq. (24) to the RMS parameter errors shown in the upper right part of Fig. 5. Then, the linear deformation error was substituted for the deformation error, since the difference was found negligible (Fig. 6).

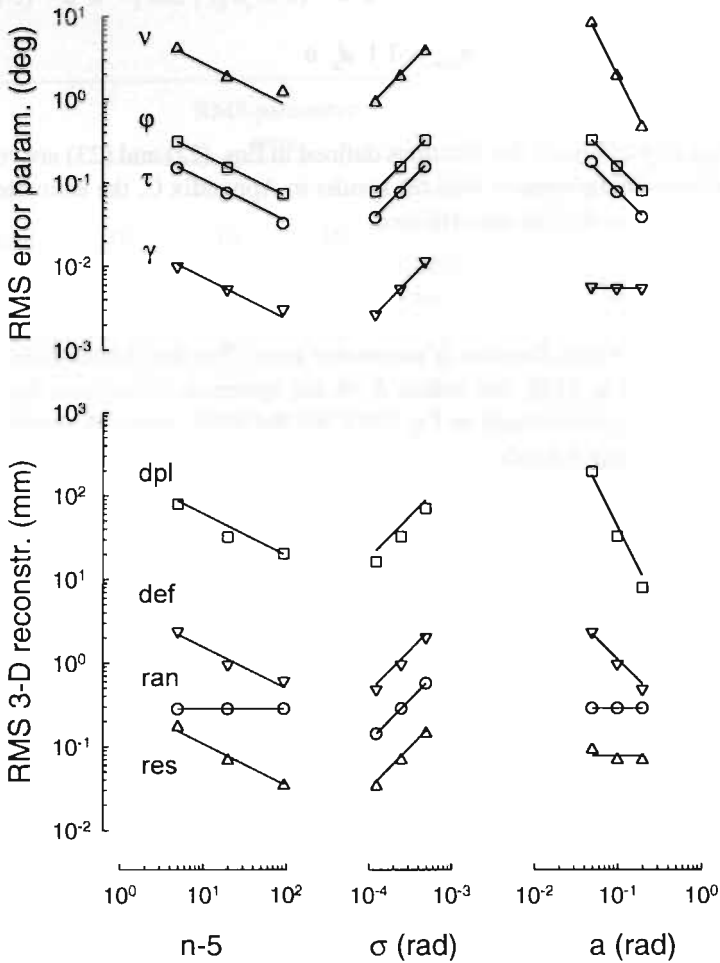


FIGURE 5. Parameter and reconstruction accuracy as a function of the number of stereo pairs n , the RMS image position noise σ , and the aperture a . Results of the ML method with computer simulations of markers attached to the surface of a sphere. Points indicate results of 100 repeated simulations per condition. Solid lines represent curves fitted to the data. *Upper panel*: RMS parameter errors ϵ_v , ϵ_ϕ , ϵ_τ , and ϵ_γ . *Lower panel*: 3-D reconstruction, RMS values ϵ_{dpl} , ϵ_{def} , ϵ_{res} , and ϵ_{ran} , respectively. The RMS values correspond to the total displacement error, the deformation error, and the residual error caused by the parameter errors, and the random error due to the image position noise.

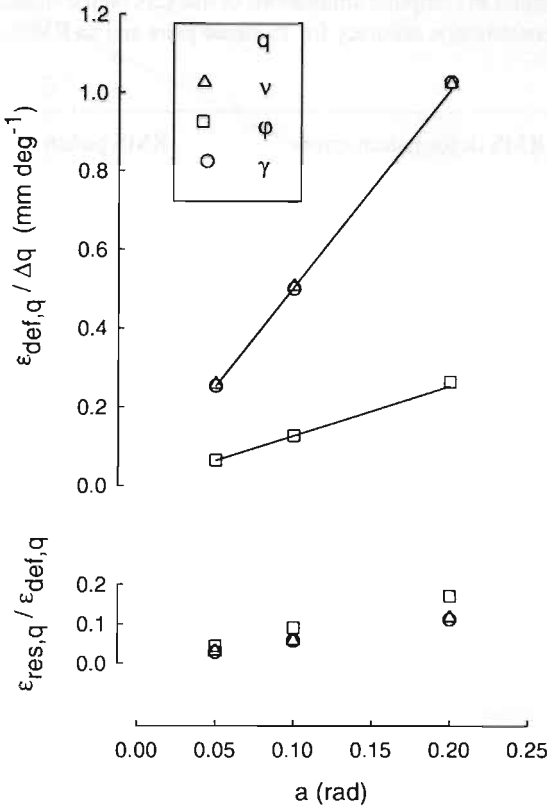


FIGURE 6. Deformation error of 3-D reconstruction induced per parameter error, as a function of the aperture a . Computer simulation of 100 markers attached to the surface of a sphere. Parameter errors were set to the values shown in the upper right part of Fig. 5. *Upper panel*: Ratio of the RMS deformation error of the 3-D reconstruction and the parameter perturbation Δq of parameter q , as a function of the aperture. Points: Simulation results. Curves: RMS linear deformation error $\epsilon_{\text{ldf},\Delta p}$ divided by parameter error. Error $\epsilon_{\text{ldf},\Delta p}$ was calculated according to mathematically derived expression $[c_q d_s a]$, with constant c_q equal to 5, 5, 1.25 [10^{-3} rad⁻¹ deg⁻¹] for parameters v , τ , and ϕ , respectively, and d_s equal to the source point distance. *Lower panel*: Ratio of the RMS residual deformation error and the RMS deformation error, as a function of the aperture.

TABLE IV. ML method in computer simulations of markers on the surface of a sphere. Parameter and reconstruction accuracy for 25 stereo pairs and an RMS position noise of 2.5×10^{-4} rad.

aperture	RMS deformation error			RMS parameter error			
	ϵ_{def}	$\epsilon_{def,\phi} / \epsilon_{def}$	$\epsilon_{def,\tau} / \epsilon_{def}$	ϵ_v	$\epsilon_\phi / \epsilon_v$	$\epsilon_\tau / \epsilon_v$	$\epsilon_\gamma / \epsilon_v$
rad	mm	%	%	deg	%	%	%
0.05	2	1	2	8	4	2	0.1
0.10	1	2	4	2	8	4	0.3
0.20	0.5	4	8	0.5	16	8	1

TABLE V. Reconstruction accuracy of ML method and LLS method in computer simulations of markers distributed over the volume of a sphere.

number of stereo pairs	image position noise		RMS deformation error of 3-D reconstruction	
	[10^{-4} rad]		[mm]	
	Gaussian, RMS value	pixel round off, amplitude	LLS method	ML method
8	0.0	1.8	5.0	0.8
10	0.0	1.8	1.0	0.6
30	0.0	1.8	0.3	0.23
10	3.5	1.8	3.5	2.0
10	7.0	1.8	8.0	4.0

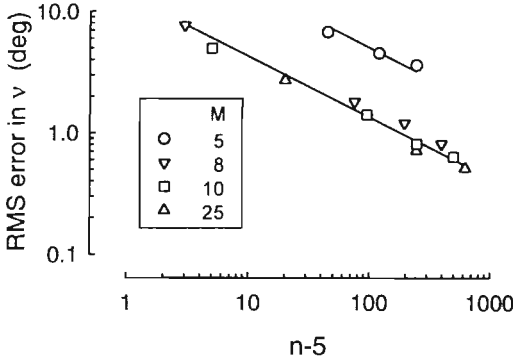


FIGURE 7. Parameter accuracy of the ML method with computer simulations of multiple frame stereo observations of LV motion. Number of markers, M , varied as indicated. Number of frames, T , set to 1, 10, 25, and 50. Results of 100 repeated simulations per condition, shown for parameter v . Points: simulation results. Curves: Fitted function $[c_f (n-5)^{-1/2}]$ with $n=M \times T$, and constant c_f .

Situation B: Markers attached to a simulated moving left ventricle

ML method, parameter accuracy. In Fig. 7 the RMS error of parameter v is presented as a function of $n-5$, the number of stereo pairs minus the number of fitted parameters. The function

$$c_f (n-5)^{-1/2} \quad n = M \times T \quad (25)$$

with c_f a constant, was LS fitted to the points in Fig. 7, using separate curves for the sets of points for $M=5$ and $M \geq 8$. The results for the parameters ϕ , τ , and γ were similar to those shown for v . With single frame data of 25 markers, the RMS parameter errors for v , ϕ , τ and γ were found equal to 2.7, 0.17, 0.08, and 0.007 deg, respectively, implying an increase of 10-30% when compared to the corresponding results with the spherical object (Table IV).

Situation C: Markers distributed over the volume of a sphere

Reconstruction accuracy of ML method and LLS method. The accuracy obtained with the LLS method in the simulations reported in [25], and with the ML method in similar simulations in the current study, respectively, is presented in Table V.

Discussion

In the present study we developed and evaluated a method to reconstruct 3-D object points from stereo images by using a maximum likelihood (ML) method to estimate the geometric parameters of a stereo camera set-up. In the used model, the cameras are represented as ideal pinhole cameras, and the image distance and the intersection of the camera axis with the image plane are assumed known. Then the geometry of the stereo camera set-up (Fig. 2) can be described by a set of 5 angles, $(\phi_L, \phi_R, \nu, \tau, \gamma)$, and the mismatch of a stereo image pair is expressed as a nonlinear function of the 5 parameters. When a sufficient number ($n \geq 5$) of stereo pairs is known, the 5 parameters can be estimated and the unknown 3-D positions of the object points can be reconstructed. The 5 parameter nonlinear model is preferred to the 8 parameter linear representation [23-25] for two reasons: 1) the model is parsimonious, since the 5 parameters correspond to the 5 degrees of freedom of the geometry, and 2) the physically interpretable parameters are appropriate to investigate the estimation method and the structure of the reconstruction error.

Assumptions. In the used model (Fig. 2) the cameras are represented as ideal pinhole cameras, and the image distance and the intersection of the camera axis with the image plane are assumed known. A real X-ray stereo camera set-up can be represented by this model when the interior calibration [22] of the imaging system is adequate. The major distortions of the imaging system can be corrected by the use of a dewarp grid [3]. The reported residual error of approximately 0.5×10^{-4} rad is small compared to the RMS noise level of 2.5×10^{-4} rad. Computer simulations showed that errors of 0.1° and 0.2° in the angles ν and τ , respectively, were induced per percentage error in the image distance, the distance from source point to dewarp grid plane. According to Eq. (24) the corresponding reconstruction errors are equal to 0.05 and 0.1 mm, respectively, which are small compared to the inevitable random reconstruction error of 0.3 mm caused by noise of an RMS level of 2.5×10^{-4} rad (see Eq. (23)). So, when the error in the image distance is smaller than 1%, the induced reconstruction error may be neglected. It can be derived that an error δ in the position of the intersection of the camera axis with the image plane causes a systematic error of the order of $u^2 \delta$ in the angle between two imaging rays, where u represents an image coordinate. With a viewing aperture of 0.1 rad, corresponding to 200 pixels, and assuming $\delta = \pm 0.005$ rad, the error introduced in the image plane is 0.02 pixel, which is small compared to the noise of 0.5 pixel. In conclusion: the systematic errors, which are introduced by inaccurate knowledge of the interior parameters of the imaging systems, in our set-up are negligible compared to the effects of noise of an RMS value of 2.5×10^{-4} rad (0.5 pixel). For substantially lower noise levels, the systematic errors may become dominating, and a more accurate interior calibration of the imaging systems is indicated.

Parameter estimation. The elevation angle difference, defined in Eq. (9), is used as a measure of stereo mismatch. The parameters are estimated by minimizing the weighted elevation angle differences of n stereo pairs in a nonlinear least squares regression procedure. When the errors in the measured image positions are assumed gaussian, maximum likelihood (ML) estimates can be obtained by using weights which are inversely proportional to the error variance in the elevation angle differences. Conventional nonlinear least squares (NLS) estimation methods, which minimize the differences of 1) the unweighted elevation angle, or 2) the elevation arc length, use a similar least squares criterion function, but nonoptimal weights.

Evaluation in computer simulations. The performance was evaluated in computer simulations of the standard stereo camera set-up defined by Eq. (19). The perpendicular camera axes of the set-up are optimal [33] with respect to a minimum 3-D reconstruction error. The major part of the evaluation was based on computer generated single frame stereo image data of markers attached to the surface of a sphere. This geometry in a stylized way represents an organ with a cavity which is observed in an X-ray stereo camera set-up. By locating the markers over the surface at random but with a minimum allowed distance between them, see Eq. (20), realistic marker distributions were obtained. Noise perturbed data were used as input for the estimation procedure, and in 100 repeated simulations the parameter bias, and the root mean square (RMS) parameter and reconstruction error were obtained. The ML method was also evaluated with multiple frame stereo image data obtained in a realistic computer simulation of markers attached to a moving left ventricle [32]. In order to compare the performance of the ML method and the 8 parameter linear least squares (LLS) method reported by Fencil and Metz [25], single frame stereo data of markers distributed over the volume of a sphere were generated.

Performance ML method and conventional NLS methods. The ML method resulted in an unbiased estimate of v , and a minimum RMS estimation and reconstruction error (Fig. 4). The estimates of the other four parameters were unbiased, and equally accurate with ML and conventional NLS. With the NLS methods, the bias in v only decreased to nonsignificant values when the aperture was increased to 0.2 rad (Fig. 4, upper panel, column d). Contrary to the NLS methods, for the ML method increasing the number of stereo pairs resulted in a significant improvement of the accuracy (lower panel, column c versus b). When the minimum number of $n=5$ stereo pairs was used, the ML-method still resulted in unbiased estimates, but the accuracy of v and τ reduced to inappropriate RMS levels of 18° and 9° , respectively. Nevertheless, in all simulations the ML method performed best, so the method is preferred to the NLS methods.

Parameter errors ML method. As is shown in Fig. 5 and Table IV, for an aperture of 0.1 rad, the error in parameter v outweighs the error in the other parameters, and dominates the reconstruction error. Obviously, the accuracy of parameter v is harmed most by the lack of perspective effect which occurs when the aperture is small. Hence,

the inherently bad conditioning of the estimation of the relative orientation in a stereo camera set-up as mentioned in the literature [25], was shown to be mainly due to the influence of a single parameter, the angle 2ν between the two camera axes. So, when the estimation accuracy is poor, improvement might be obtained by estimating parameter ν independently, e.g. by fixing it at the intended value of 45° .

As expected, the RMS parameter errors were proportional to the RMS position noise, and inversely proportional to the square root of the number of stereo pairs minus the number of fitted parameters (Fig. 5, solid lines). Furthermore, in agreement with the derivation in Appendix C, the parameter estimates were found to be uncorrelated, and the RMS parameter errors were inversely proportional to the aperture for τ , ϕ_L , and ϕ_R , inversely proportional to the squared aperture for ν , and unaffected by the aperture for γ .

Reconstruction errors ML method. When the position of the object with respect to the camera system is relevant, as for instance in robotics [22], the total displacement error defined in Eq. (14) is of interest. However, when measuring deformation, the uniform translation, rotation and scaling of the object are not of interest, and the deformation error defined in Eq. (17) is the only relevant part of the reconstruction error. In all simulations of the current study, the residual error defined in Eq. (18) was very small (Fig. 6), and the deformation error was approximately equal to the linear deformation error (difference $< 2\%$). The error in parameter ν dominated the reconstruction error (Table IV), and in agreement with Eq. (24), the RMS deformation error, as a function of the number of stereo pairs and the noise level, behaved like the RMS error in ν , and was inversely proportional to the aperture. For the standard values $n=25$, $\sigma=2.5 \times 10^{-4}$ rad, and $\alpha=0.1$ rad, the RMS deformation error was equal to 1 mm, and was mainly due to the RMS parameter error of 2° in ν (Table IV).

Improving reconstruction accuracy by increasing n . As shown in Fig. 5, the parameter and reconstruction accuracy can be improved by increasing the aperture, the number of stereo constraints, or the imaging accuracy. Often the maximum aperture and the imaging accuracy are fixed properties of the available equipment, and improvement can only be achieved by increasing the number of stereo pairs. The extent of the improvement, is limited by the inevitable random reconstruction error (see Section ‘3-D Reconstruction error’), whose magnitude depends on the noise level, but is not affected by the accuracy of the parameter estimation (Fig. 5). On the basis of Eqs. (22) and (23), and the coefficients in Table III, the ratio $r_{dr} = \varepsilon_{def} / \varepsilon_{ran}$ can be defined as a function of the aperture α and the number of stereo pairs n ,

$$r_{dr} = 1.8 \times (n-5)^{-1/2} \alpha^{-1} \quad (26)$$

and with the aperture set to 0.1 rad, ratio values $r_{dr}=1$, 0.5, and 0.1 are obtained for $n=300$, 1300, and 30,000, respectively. With single frame observations these numbers

may not seem practical, but with multiple frame data a number of 1250 stereo pairs can very well be obtained by observing $M=25$ markers in $T=50$ frames. For $M=25$ and $\alpha=0.1$ rad, application of Eq. (26) with $M \times T$ substituted for n , results in a reduction of r_{dr} from 4 to 0.5 when increasing the number of frames from $T=1$ to 50.

ML method with simulated LV motion. When more frames are used, the number of stereo pairs increases proportionally. The results of simulations with multiple frame observations (Fig. 7), for $M \geq 8$ show that the RMS parameter error varies proportional to $(M \times T \cdot 5)^{-1/2}$. Hence, an equivalent improvement of accuracy is obtained by increasing the number of markers, M , or by increasing the number of frames, T , provided the markers move considerably, as is the case with myocardial markers. For $M=5$, an increased RMS level was found, which is probably due to a decrease of effective aperture. When a set of projected markers does not completely span the area corresponding to the object aperture a , the effective aperture is smaller than a . According to Eq. (22) and the coefficients in Table III, a decrease of aperture implies a decrease of estimation accuracy. With low M , distributions of markers which do not span the complete object aperture are more likely to occur.

The increase of 10-30% of the RMS parameter error, which was observed when the spherical object was replaced by a simulated left ventricle, is probably also due to a decrease of effective aperture. When markers are distributed over an objects volume instead of its outer surface, the spread of the marker distribution on average will be smaller, and, as a result, the effective aperture will be decreased.

Reconstruction accuracy of ML and LLS method. Comparison of the results of the ML method with those of the 8 parameter linear least squares (LLS) method [25], shows that the RMS deformation error for the ML method ranges from 50% to 80% of the levels reported for the LLS method (Table V). For $n=8$, which is the minimum number of stereo correspondences needed for application of the 8 parameter LLS method, the resulting percentage was even less than 20%. The better performance of the ML method is likely to be due to the smaller number of parameters, and the optimality of the method with regards to the RMS error as well as to the bias of, in particular, the estimate of parameter v .

Conclusions

Application of the maximum likelihood (ML) method to estimate the relative orientation in a stereo camera set-up with perpendicular camera axes and an aperture of approximately $\alpha=0.1$ rad, results in unbiased and uncorrelated estimates of the five angular parameters of the used nonlinear model. The minimum number of stereo correspondences, which are needed as constraints, is equal to 5. In computer simulations with $n=25$ markers positioned at the surface of a sphere of 70 mm diameter, and the image position noise set to a root mean square (RMS) value $\sigma=2.5 \times 10^{-4}$ rad (0.5 pixel

in a 256×256 pixel image matrix), the RMS deformation error of the reconstruction was found to be equal to 1 mm.

Conventional nonlinear least squares methods based on minimizing the elevation angle or the elevation arc differences, resulted in a significantly biased estimate of parameter v , and in larger reconstruction errors. Contrary to the ML method, the errors did not decrease when the number of stereo pairs was increased.

For markers attached to a sphere of 70 mm diameter and an object aperture of 0.1 rad, it was derived that a reconstruction error of 0.5 mm was induced per error of 1° in parameters v and τ , and a reconstruction error of 0.13 mm for a 1° error in parameters ϕ_L , or ϕ_R . An error in parameter γ did not induce a reconstruction error at all.

Parameter v , which represents the angle between the two camera axes, was the least accurately estimated and dominated the reconstruction error. In the simulations specified above, the RMS error in v was 2° , and it was shown to vary proportional to $[(n-5)^{-1/2} \sigma \alpha^{-2}]$. The resulting RMS deformation error varied in a similar way, but proportional to α^{-1} .

Application of the ML method to stereo data obtained in computer simulations with the spherical object replaced by a moving left ventricle, for single frame data resulted in similar parameter and reconstruction accuracies. With multiple frame data, using the stereo correspondences of M markers in T frames as constraints, the accuracy was shown to vary proportional to $(M \times T - 5)^{-1/2}$. So, similar improvements of the accuracy were obtained, whether the set of stereo pairs was extended by using more markers, or by using more frames.

The ratio r_d of the RMS value of the systematic deformation error due to the parameter errors, and the inevitable random reconstruction error due to the image position noise, was shown to be equal to $[1.8 (n-5)^{-1/2} \alpha^{-1}]$. Hence, if improvement of the accuracy is considered irrelevant below $r_d=1$, for apertures $\alpha=0.05, 0.1, 0.2$, and any noise level, it is irrelevant to use larger numbers of stereo correspondences than $n=1200, 300$, and 80, respectively.

Compared with the results of 3-D reconstruction by using the 8 parameter linear least squares (LLS) method [25], the ML method resulted in RMS deformation errors ranging from 50-80% of the levels reported for the LLS method.

Appendix A

Maximum Likelihood estimates

Let the measured stereo image coordinate vector u be defined

$$u = (u_L, v_L, u_R, v_R)^T \quad (27)$$

and assume that the coordinates are perturbed by additive gaussian noise, which is independently and identically distributed for the four coordinates, then

$$\mathbf{u} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}; \quad \boldsymbol{\mu} = E(\mathbf{u}); \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 I) \quad (28)$$

with 4×1 vectors $\boldsymbol{\mu}, \boldsymbol{\varepsilon}$ representing the true value and the error, respectively, E the mathematical expectation, N indicating the multivariate normal distribution, $\mathbf{0}$ a 4×1 vector of zeros, σ^2 the noise variance, and I the 4×4 unit matrix. The model defined by Eq. (28) is invariant under orthogonal linear transformation, so for the image coordinates $\mathbf{u}' = (u'_L, v'_L, u'_R, v'_R)$, defined in Eq. (7), a model similar to Eq. (28) holds with $\mathbf{u}, \boldsymbol{\mu}$, and $\boldsymbol{\varepsilon}$ replaced by $\mathbf{u}', \boldsymbol{\mu}'$, and $\boldsymbol{\varepsilon}'$, respectively. Then, for the elevation angle β_L (Fig. 2, middle panel) defined in Eq. (4), it holds

$$\beta_L = \arctan(v'_L / r'_L) \quad (29)$$

$$r'_L = |P'_L P''_L| = \sin \omega_L - u'_L \cos \omega_L \quad u'_L = \mu'_1 + \varepsilon'_1 \quad v'_L = \mu'_2 + \varepsilon'_2$$

with points P'_L, P''_L defined in Fig. 2, and μ'_1, μ'_2 and $\varepsilon'_1, \varepsilon'_2$ the first two elements of $\boldsymbol{\mu}'$ and $\boldsymbol{\varepsilon}'$, respectively. When the aperture is small, $a \approx 0.1$, and σ is of the order of $0.01a$, angle β_L may be approximated

$$\beta_L \approx E(\beta_L) + \varepsilon'_2 / \rho_L \quad \rho_L = E(r'_L) \quad (30)$$

Using a similar approximation for β_R , and noting that the expected value of the stereo mismatch $\Delta \beta$, defined in Eq. (9), is zero, for the mismatch of the i -th stereo pair it holds

$$\Delta \beta_i = \rho_{L,i}^{-1} \varepsilon_{2,i}' - \rho_{R,i}^{-1} \varepsilon_{4,i}' \rightarrow \Delta \beta_i \sim N(0, (\rho_{L,i}^{-2} + \rho_{R,i}^{-2}) \sigma^2) \quad (31)$$

with $r_{L,i}$, the distance defined in Eq. (29), $\rho_{L,i}$ the expected value of $r_{L,i}$, $\rho_{R,i}$ similarly defined for the right image, and $\varepsilon_{2,i}', \varepsilon_{4,i}'$ components of the error vector $\boldsymbol{\varepsilon}'$ for the i -th stereo pair, $i=1, 2, \dots, n$.

Eq. (31) implies that the measured elevation angle differences $\Delta \beta_i$, are not equally reliable for all i . In least squares optimization the reliability differences are taken into account by using weights which are inversely proportional to the error variance [34]. Then, for n stereo pairs we have the weighted least squares estimate

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} \mid \sum_{i=1}^n \Delta \beta_i^2 w_i^2 \text{ minimum} \quad w_i^2 = (\rho_{L,i}^{-2} + \rho_{R,i}^{-2})^{-1} \quad (32)$$

which is the ML estimate when the errors are gaussian [34].

In our implementation, $r_{L,i}, r_{R,i}$ is substituted in w_i^2 for the unknown $\rho_{L,i}, \rho_{R,i}$. With σ of the order of $0.01a$, the resulting change of the weight is negligible.

Appendix B

Decomposition of uniform linear deformation

By the polar decomposition technique [36,37], a linear deformation matrix F can be decomposed into a rotation and a pure deformation, and it holds

$$F = RD \quad (33)$$

with R a 3×3 orthogonal rotation matrix and D a 3×3 symmetrical stretch matrix. The deformation D can be decomposed [38] into an incompressible deformation represented by matrix G , and a scale factor s , so we have

$$\begin{aligned} F &= sRG \\ s &= (\text{trace}(D))^{1/3} \quad G = s^{-1}D \end{aligned} \quad (34)$$

The matrices R , and D , and the scalar s are readily obtained from the Singular Value Decomposition [39] of matrix F

$$\begin{aligned} F &= A \Lambda^{1/2} B^T \\ R &= AB^T \quad D = B \Lambda^{1/2} B^T \quad s = (\text{trace}(\Lambda^{1/2}))^{1/3} \end{aligned} \quad (35)$$

where A and B are 3×3 matrices whose columns correspond to the eigenvectors of matrices FF^T , and F^TF , respectively, and matrix $\Lambda^{1/2}$ is a 3×3 diagonal matrix of singular values [39].

Appendix C

Linear approximation of the parameter estimation

Assume the standard stereo camera set-up defined in Eq. (19), and let $\Delta\theta$ be a vector of deviations from the true parameter value

$$\Delta\theta = (\Delta\varphi_L, \Delta\varphi_R, \Delta v, \Delta\tau, \Delta\gamma)^T \quad (36)$$

then, on the basis of Eqs. (4), (6), and (7), the stereo constraint corresponding to Eq. (9) can be written

$$\arctan(\dot{v}_L / (\sin \eta_L - \dot{u}_L \cos \eta_L)) - \arctan(\dot{v}_R / (\sin \eta_R - \dot{u}_R \cos \eta_R)) - \Delta \gamma \approx 0 \quad (37)$$

$$\eta_L = \pi/4 - \Delta v + \Delta \tau \quad \eta_R = 3\pi/4 + \Delta v + \Delta \tau$$

For an aperture $a \approx 0.1$, the constraint can be approximated

$$\begin{aligned} & \dot{v}_L - \dot{v}_R + \dot{v}_L \dot{u}_R + \dot{u}_L \dot{v}_R - (\dot{v}_L - \dot{v}_R - \dot{v}_L \dot{u}_R - \dot{u}_L \dot{v}_R) \Delta v + \\ & - (\dot{v}_L + \dot{v}_R) \Delta \tau - \Delta \gamma / \sqrt{2} \approx 0 \end{aligned} \quad (38)$$

For small angles $\Delta \varphi_L$ and $\Delta \varphi_R$, substituting Eq. (28), Appendix A, and ignoring the higher order terms, it approximately holds

$$\begin{aligned} \dot{u}_L &= \mu_1 + \varepsilon_1 + \mu_2 \Delta \varphi_L & \text{and} & & \dot{u}_R &= \mu_3 + \varepsilon_3 + \mu_4 \Delta \varphi_R \\ \dot{v}_L &= -\mu_1 \Delta \varphi_L + \mu_2 + \varepsilon_2 & & & \dot{v}_R &= -\mu_3 \Delta \varphi_R + \mu_4 + \varepsilon_4 \end{aligned} \quad (39)$$

with $\varepsilon_1, \dots, \varepsilon_4$, and μ_1, \dots, μ_4 elements of vectors ε and μ defined in Eq. (28). Substitution of Eq. (39) in Eq. (38) and ignoring the higher order terms, results in

$$\begin{aligned} & \mu_4 - \mu_2 + \mu_2 \mu_3 + \mu_1 \mu_4 \approx \mu_1 \Delta \varphi_L - \mu_3 \Delta \varphi_R + \\ & + (\mu_2 - \mu_4 - \mu_2 \mu_3 - \mu_1 \mu_4) \Delta v + (\mu_2 + \mu_4) \Delta \tau + \Delta \gamma / \sqrt{2} + \varepsilon_2 - \varepsilon_4 \end{aligned} \quad (40)$$

which for n stereo constraints is equivalent with the linear regression model

$$\mathbf{g} = H \Delta \theta + \xi \quad \xi \sim N(0, 2\sigma^2 I) \quad (41)$$

with \mathbf{g} an $n \times 1$ vector corresponding to the left-hand side of Eq. (40), H an $n \times 5$ matrix whose columns correspond to the coefficients of the right-hand side, $\Delta \theta$ the parameter deviation vector defined in Eq. (36), and ξ an $n \times 1$ vector of independent gaussian errors corresponding to the error term $\varepsilon_2 - \varepsilon_4$.

For a configuration of object points with spherical symmetry, e.g. a homogeneous distribution of markers at the surface of a sphere, it can be verified that the columns of matrix H are approximately orthogonal, so the parameter estimates are uncorrelated. Furthermore, according to the definition of the aperture a in the Section 'Performance

evaluation', a change of a results in a proportional change of μ_i , $i=1,\dots,4$, while the error terms ε_i , $i=1,\dots,4$, remain the same. So, the columns of matrix H which correspond to $\Delta\phi_L$, $\Delta\phi_R$, and $\Delta\tau$ are proportional to a . The column corresponding to Δv , however, can be shown to be approximately proportional to the product of two 3-D coordinates, and hence to a^2 . For $\Delta\gamma$, the corresponding column is independent of a . Then, for the covariance matrix [35] of the parameter estimates, it holds

$$V(\Delta\theta) = (H^T H)^{-1} 2\sigma^2 = \text{diag}(d_1 a^{-2}, d_2 a^{-2}, d_3 a^{-4}, d_4 a^{-2}, d_5) 2\sigma^2 \quad (42)$$

where $\text{diag}(\dots)$ represents a diagonal matrix, with the diagonal elements indicated between the brackets. The value of the constants d_1 to d_5 depends on the particular configuration of the markers.

Appendix D

Reconstruction error induced per parameter

The upper panel of Fig. 8 shows, in a view from above, the standard stereo camera set-up with an object point P and the associated pair of intersecting stereo rays $(S_L P_L, S_R P_R)$. The effect of an error Δv is introduced by rotating the left and right image ray bundles towards each other, over an angle Δv around perpendicular axes in S_L and S_R . The rotated rays $(S_L P_L', S_R P_R')$ in general do not intersect (Fig. 8, lower panel) and the location of the reconstructed 3-D point \hat{P} is approximated by the midpoint between the crossing rays, as indicated in the inset. For an aperture $a \approx 0.1$ rad, the displacement error is approximately

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} - \begin{bmatrix} x \\ y \\ z \end{bmatrix} \approx \begin{bmatrix} 0 \\ -2\Delta v y - \Delta v \\ -\Delta v z \end{bmatrix} \quad (43)$$

with $(\hat{x}, \hat{y}, \hat{z})^T$ and $(x, y, z)^T$ the position vectors of points P and \hat{P} , respectively. Decomposition of the displacement according to Eqs. (15) and (16), results in

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} \approx \begin{bmatrix} 0 \\ -\Delta v \\ 0 \end{bmatrix} + (1 - \Delta v) \begin{bmatrix} 1 + \Delta v & 0 & 0 \\ 0 & 1 - \Delta v & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (44)$$

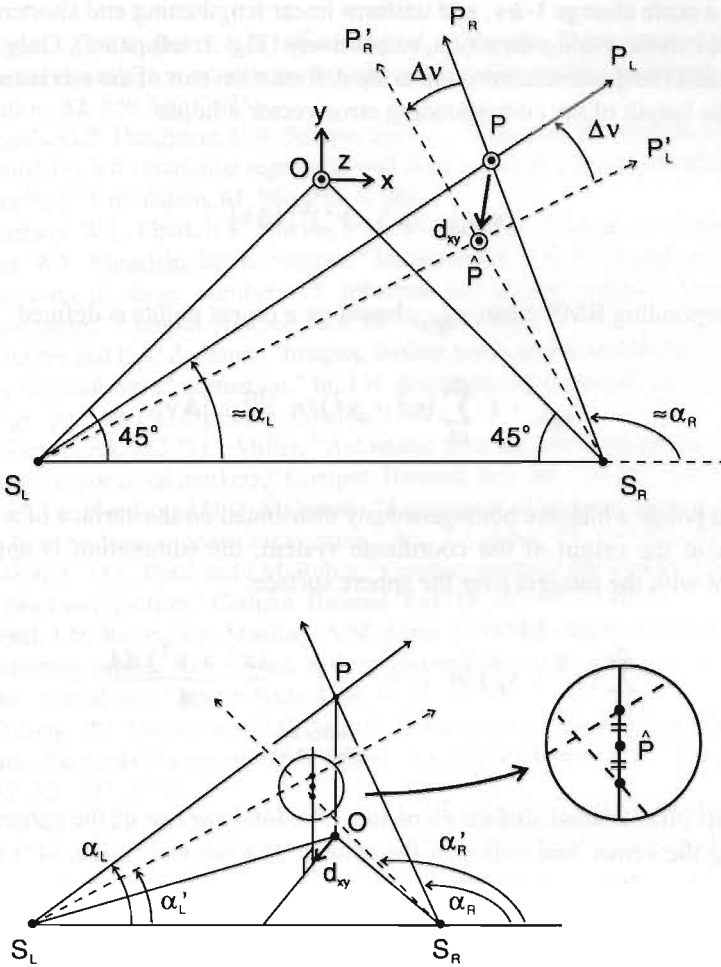


FIGURE 8. Mathematical derivation of the reconstruction error induced by a parameter error. *Upper panel:* Orthogonal view from above. Standard stereo camera set-up with source points S_L and S_R , and xyz -coordinate system originated in O . With object point P corresponds the exact image ray pair $(S_L P_L, S_R P_R)$, and α_L and α_R are the angles of the image rays with baseline $S_L S_R$. For aperture $a=0.1$ rad, the angles α_L and α_R approximately hold also for the projection of the rays in the xy -plane. Ray pair $(S_L P'_L, S_R P'_R)$ results when parameter v is perturbed by an error Δv , and point \hat{P} is the corresponding erroneous reconstruction of P . Vector d_{xy} represents the displacement error in the xy -plane. *Lower panel:* Approximate displacement error in z -direction corresponding to the displacement d_{xy} .

Hence, an error $\Delta \mathbf{v}$ results in a translation of the object over a distance $\Delta \mathbf{v}$ towards the baseline, a scale change $1 - \Delta \mathbf{v}$, and uniform linear lengthening and shortening proportional to $\Delta \mathbf{v}$ in the x and y direction, respectively (Fig. 3, left panel). Only the last two displacement components contribute to the deformation part of the reconstruction error, and for the length of the corresponding error vector it holds

$$|e_{ldf,v}| = (x^2 + y^2)^{1/2} |\Delta \mathbf{v}| \quad (45)$$

The corresponding RMS error, $\epsilon_{ldf,v}$, based on n object points is defined

$$\epsilon_{ldf,v} = \left(\sum_{i=1}^n (x_i^2 + y_i^2) / n \right)^{1/2} |\Delta \mathbf{v}| \quad (47)$$

For object points which are homogeneously distributed on the surface of a sphere with its centre at the origin of the coordinate system, the summation is approximately equivalent with the integral over the sphere surface

$$\sum_{i=1}^n (x_i^2 + y_i^2) / n \approx \int_{\substack{\text{surface of} \\ \text{sphere } (0,R)}} \frac{(x^2 + y^2) dA}{A} \quad (48)$$

with dA an infinitesimal surface element, A the total surface of the sphere, and $(0,R)$ indicating the centre and radius of the sphere. The integral in Eq. (47) evaluates to $(2/3)R^2$, so Eq. (46) can be written

$$\epsilon_{ldf,v} = (2/3)^{1/2} R |\Delta \mathbf{v}| \quad (49)$$

In a similar way, the approximate displacements caused by errors in the other parameters are derived. Errors $\Delta \tau$, and $\Delta \phi$ in parameter τ , and ϕ_L or ϕ_R , respectively, also result in proportional deformation errors corresponding to lengthening and shortening in two orthogonal directions. For the corresponding RMS linear deformation errors $\epsilon_{ldf,\tau}$, and $\epsilon_{ldf,\phi}$, it can be shown that an expression similar to Eq. (48) holds with $\Delta \mathbf{v}$ replaced by $\Delta \tau$, and $\Delta \phi/4$, respectively. An error $\Delta \gamma$ is introduced by rotating the ray bundles of the left and right camera over angles $-\Delta \gamma/2$ and $+\Delta \gamma/2$, to the right around baseline vector $\overline{S_L S_R}$. Then no reconstruction error results, since the reconstructed 3-D points calculated according to Eq. (5) exactly coincide with the original object points.

References

- [1] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers," *Circulation*, **52**, 859-867 (1975).
- [2] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard," *Circulation*, **61**, 966-972, (1980).
- [3] J.B. Garrison, W.L. Ebert, R.E. Jenkins, S.M. Yionoulis, H. Malcom, G.A. Heyler, A.A. Shoukas, W.L. Maughan, and K. Sagawa, "Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms," *Comput. Biomed. Res.* **15**, 76-96 (1982).
- [4] W.C. Hunter, and E.A. Zerhouni, "Imaging distinct points in left ventricular myocardium to study regional wall deformation," In: J.H. Anderson (ed.), *Innovations in diagnostic radiology*, pp. 169-190 (Springer, London, 1989).
- [5] M.A. Niczyporuk and D.C. Miller, "Automatic tracking and digitization of multiple radiopaque myocardial markers," *Comput. Biomed. Res.* **24**, 129-142 (1991).
- [6] Y. Kong, J.J. Morris and H.D. McIntosh, "Assessment of regional myocardial performance from biplane coronary angiograms," *Am. J. Cardiol.* **27**, 529-537 (1971).
- [7] S.A. MacKay, M.L. Potel and J.M. Rubin, "Graphics methods for tracking three-dimensional heart wall motion," *Comput. Biomed. Res.* **15**, 455-473 (1982).
- [8] M.J. Potel, J.M. Rubin, S.A. MacKay, A.M. Aisen, J. Al-Sadir and R.E. Sayre, "Methods for evaluating cardiac wall motion in three dimensions using bifurcation points of the coronary arterial tree," *Invest. Radiol.* **18**, 47-57 (1983).
- [9] A.A. Young, P.J. Hunter and B.H. Smaill, "Estimation of epicardial strain using the motions of coronary bifurcations in biplane cineangiography," *IEEE Trans. Biomed. Eng.* **39**, 526-531 (1992).
- [10] P.A. Chevalier, J.R. Rodarte, and L.D. Harris, "Regional lung expansion at total lung capacity in intact vs. excised canine lungs," *J. Appl. Physiol.: Respirat. Environ. Exercise. Physiol.* **45**, 363-369 (1978).
- [11] J.H. Wei, E.A. Hoffman, E.L. Ritman, and E.H. Wood, "Cardiogenic motion of the right lung parenchyma in anesthetized intact dogs," *J. Appl. Physiol.* **58**, 384-391 (1985).
- [12] B.A. Simon and W. Mitzner, "Redistribution of pulmonary ventilation during high-frequency," *Fed. Proc.* **46**, 1424 (1987).
- [13] J.L. Péan, C.J. Chuong, M. Ramanathan and R.L. Johnson, "Regional deformation of the canine diaphragm," *J. Appl. Physiol.* **71**, 1581-1588 (1991).
- [14] S.S. Kramer, J.H. Anderson, J.D. Strandberg and M.W. Donner, "A permanent radiopaque marker technique for the study of pharyngeal swallowing in dogs," *Dysphagia*, **1**, 163-167 (1987).
- [15] K.M. Hiimeae, and A.W. Crompton, "Mastication, food transport, and swallowing". In: M. Hildebrand, D.M. Bramble, K.F. Liem, D.B. Wake (eds.), *Functional vertebrate morphology*, pp. 262-290 (Harvard Univ. Press, Cambridge, MA, 1985).
- [16] J.B. Palmer, E. Tanaka, and A.A. Siebens, "Motions of the posterior pharyngeal wall in swallowing," *Laryngoscope*, **98**, 414-417 (1988).
- [17] N.A. Borghese, and G. Ferrigno, "An algorithm for 3-D automatic movement detection by means of standard TV cameras," *IEEE Trans. Biomed. Eng.* **37**, 1221-1225 (1990).

- [18] G.W.M.P. Peters, *Tools for the measurement of stress and strain fields in soft tissue*, (Ph.D thesis, University of Limburg, Maastricht, The Netherlands, 1987).
- [19] J. Downs, H.R. Halperin, J. Humphrey, and F. Yin, "An Improved Video-Based Computer Tracking System for Soft Biomaterials Testing," *IEEE Trans. Biomed. Eng.* **37**, 903-907 (1990).
- [20] G.A. Wood and R.N. Marshall, "The accuracy of DLT extrapolation in three-dimensional film analysis," *J. Biomech.* **19**, 781-785 (1986).
- [21] G.T. Daughters, W.J. Sanders, D.C. Miller, A. Schwarzkopf, W.M. Mead, and N.B. Ingels, "A comparison of two analytical systems for 3-D reconstruction from biplane videoradiograms," *Proc. Conf. on Comput. in Cardiol.* Washington 23-25 sept. 1988, 79-82 (1989).
- [22] B.P.K. Horn, *Robot Vision*, (McGraw-Hill, New York, 1986).
- [23] H.C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, **293**, 133-135 (1981).
- [24] C.E. Metz and L.E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory," *Med. Phys.* **16**, 45-51 (1989).
- [25] L.E. Fencil, and C.E. Metz, "Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation, *Med. Phys.* **17**, 951-961 (1990).
- [26] J. Weng, T.S. Huang and N. Ahuja, "Motion and structure from two perspective views: algorithms, error analysis, and error estimation," *IEEE Trans. Patt. An. Mach. Intell.* **11**, 451-476, (1989).
- [27] J.W. Roach and J.K. Aggarwal, "Determining the movement of objects from a sequence of images," *IEEE Trans. Patt. An. Mach. Intell.* **2**, 554-562 (1980).
- [28] A. Mitiche and J.K. Aggarwal, "A computational analysis of time-varying images," In: T.Y. Young and K-S Fu (eds.), *Handbook of pattern recognition and image processing*, pp. 311-332 (Academic, London, 1986).
- [29] K.L. Lam, and R.K. Ten Haken, "Improvement of precision in spatial localization of radiopaque markers using the two-film technique," *Med. Phys.* **18**, 1126-1131 (1991).
- [30] R.O. Duda, and P.E. Hart, *Pattern classification and scene analysis*, (Wiley, New York, 1973).
- [31] Numerical Algorithms Group, *NAG Fortran Library Manual*, Mark 13, (NAG, Oxford (UK), 1988).
- [32] A.M.M. Muijtjens, J.M.A. Roos, T. Arts, A. Hasman and R.S. Reneman, "Tracking markers with missing data by iterative lower rank approximation," (Chpt. 4 of this thesis).
- [33] S.K. Ghosh, *Analytical Photogrammetry*, (Pergamon, New York, 1979).
- [34] Y. Bard, *Nonlinear Parameter Estimation*, (Academic Press, New York, 1974).
- [35] G.A.F. Seber and C.J. Wild, *Nonlinear Regression*, (Wiley, New York, 1989).
- [36] G.D. Meier, M.C. Ziskin, W.P. Santamore and A.A. Bove, "Kinematics of the beating heart," *Trans. Biomed. Eng.* **27**, 319-329 (1980).
- [37] S.C. Hunter, *Mechanics of continuous media*, (Horwood, Chichester, 1983).
- [38] L.E. Malvern, *Introduction to the mechanics of a continuous medium*, (Prentice-Hall, Englewood Cliffs N.J., 1969).
- [39] G.W. Stewart, *Introduction to Matrix Computations*, (Academic, London, 1973).

Establishing stereo correspondence for marker tracks

Arno M.M. Muijtjens, Jef M.A. Roos, Theo Arts*, Arie Hasman, and Robert S. Reneman**

Departments of Medical Informatics, Biophysics, and Physiology**,
Cardiovascular Research Institute Maastricht
University of Limburg, The Netherlands*

Submitted

Abstract : Motion and deformation of the cardiac wall may be measured by following the motion of implanted radiopaque markers in 3-D, using two X-ray cameras simultaneously. Calibration of the position measurement system can be obtained from the images during the measurements, provided that at least 5 stereo pairs of marker images are available for the estimation of the 5 stereo set-up parameters. Image positions of a marker obtained in two views obey the epipolar constraint, i.e. they lie in the plane defined by the camera source points and the 3-D position of the marker. Knowledge of the camera set-up geometry is necessary to apply the epipolar constraint. Ambiguities arise when several markers lie in one epipolar plane. We developed the Iterative Track Matching (ITM) method to obtain a set of stereo pairs of marker images allowing the calibration of the stereo camera set-up. Stereo correspondences are established automatically when large numbers (10-100) of identically shaped markers are observed in an uncalibrated X-ray stereo camera set-up. It is assumed that the sequences of marker image positions measured in time per view, the so-called marker tracks, are available for at least 5 markers. Furthermore, the angle between the camera axes is assumed to be known within $\pm 20^\circ$. Ambiguities are solved by identifying corresponding tracks instead of single points. Singular Value Decomposition of the set of tracks is used to reduce noise influence and computation time. In a computer simulation 48 markers were attached to a moving heart observed in a stereo camera set-up with approximately perpendicular camera axes. The object was imaged on video (256×256 pixels, 50 frames), with an object aperture of 0.1 rad and a root mean square (RMS) position noise level of 0.5 pixel. Per view 30% of the tracks were removed at random, so that 24 stereo pairs of tracks were present, in each view supplemented by 10 tracks without a match in the other view. In 100 repeated simulations, the ITM method rendered on the average 23 stereo track pairs, 0.5% of which were false. The method also produced accurate estimates of the angular parameters of the stereo camera set-up (RMS error 0.02° to 1°) which can be used as initial values in the calibration.

Introduction

In physiological experiments regional motion and deformation of the cardiac wall may be assessed by following implanted radiopaque markers [1-5] or bifurcation points of the coronary artery tree [6-9]. The 3-D positions of the markers can be reconstructed from stereo pairs of projections, obtained by simultaneously operated X-ray cameras. Comparable set-ups were described for measuring deformation of the lung [10-12], the diaphragm [13], and the pharynx [14-16]. A similar analysis was applied to the projections of optical markers to measure human motion [17] or deformation of biological tissues [18,19].

When following attached markers or feature points, the identification of the same object point in different frames is a major problem, the so-called correspondence problem [20,21]. With stereo observations, correspondence in time as well as between the two views must be established. In some studies [3,22] temporal and stereo constraints were combined when following markers or features in 3-D. The current study focuses on the problem of finding stereo correspondence between views when the correspondences in time per view are already known.

In the computer vision literature [23-25] two major approaches to establish stereo correspondences are distinguished: area- and feature-based matching techniques. In area-based matching, correlation-based techniques are used to find subimages with similar intensity patterns. In feature-based matching the basic units to be matched are intensity anomaly features, e.g. edges or image patches, or semantic features with known physical properties and/or spatial geometry, e.g. corners or peaks of an object.

When using the images of identical radiopaque markers to serve as semantic features, the intensity patterns can not be used for identification, and matches must be solely based on the positional information. Smoothness of disparity, the relative displacement of the image of a feature point in two views, is often used as a constraint to reduce ambiguity in establishing stereo correspondence [26-29]. When radiopaque markers are observed in an X-ray stereo camera set-up, however, the disparities of neighbouring marker images may differ largely because the corresponding markers may be distant in 3-D.

Marker images which correspond to the same marker obey the epipolar constraint [23]. Representing the left camera by source point S_L , through which all image rays pass, and image plane I_L (Fig. 1, upper panel), the right camera similarly by point S_R and plane I_R , two marker images, P_L and P_R , which correspond to the same marker, lie in a plane defined by the source points and the location P of the marker. Ambiguity arises when several markers lie in the same epipolar plane (Fig. 1, middle panel). Earlier we have shown [32] that in each view a large number of identical markers (up to 100) can be reliably followed in time, resulting in two sets of marker tracks, each track consisting of a sequence of marker image positions as observed in consecutive frames. By matching marker tracks instead of single marker images the ambiguities can be solved (Fig. 1, lower panel). In practice, marker image positions are perturbed with noise and the epipolar constraint must be relaxed: the distance between image rays $S_L P_L$ and $S_R P_R$,

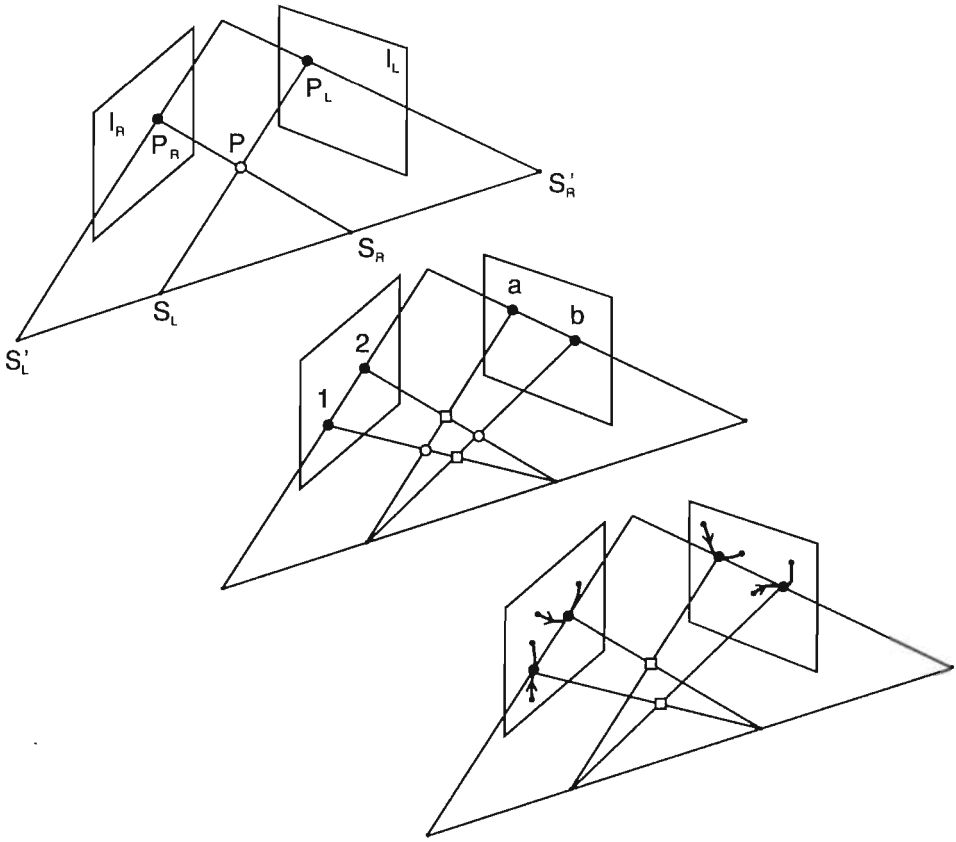


FIGURE 1. Solving ambiguities in stereo correspondence by matching marker tracks. *Upper panel:* A marker positioned at point P is simultaneously observed by two cameras, L and R . The cameras are represented by the image planes I_L and I_R , and the source points S_L and S_R through which all image rays pass. Points S_L' and S_R' represent the intersections of the baseline $S_L S_R$ with the two image planes. When marker images at positions P_L and P_R in image planes I_L and I_R , respectively, correspond to the same marker, the image rays $S_L P_L$ and $S_R P_R$ intersect in point P , the position of the marker. Hence, the points P_L and P_R lie in the indicated plane through the source points, a so-called epipolar plane. *Middle panel:* When more marker images (filled circles) lie in one epipolar plane, ambiguity arises. Two equally valid solutions exist for the stereo correspondence: stereo pairs (1,a) and (2,b), corresponding to physical markers located at the circles, and pairs (1,b) and (2,a) associated with the squares. *Lower panel:* The ambiguity is solved by matching tracks instead of single marker image positions. Curved lines indicate marker tracks. The direction of motion (arrows) and the observed positions in three consecutive frames are indicated (dots). Only the pairs of tracks associated with the squares satisfy the epipolar constraint in all three frames, and thus represent the correct stereo pairs.

or some other measure of the stereo mismatch, should be small when images P_L and P_R correspond to the same object point [4]. In order to calculate the mismatch for a tentative stereo pair, the geometry of the stereo camera set-up [30] should be known. On the other hand, if correct stereo pairs are known, the corresponding epipolar constraints can be used to calibrate the stereo camera set-up.

In recently developed systems for measuring 3-D heart motion by following radiopaque markers [3-5,7], the geometry of the stereo camera set-up is calibrated in a separate procedure by using a calibration object with a number of markers at known positions. Furthermore, limited numbers of markers (14-22) are used and stereo correspondence is established in semi-automatic procedures by applying the epipolar constraint in the known camera set-up geometry. Calibration methods which are based on the marker image position data obtained during the measurements have also been presented [31,33,35]. Then, the stereo correspondence of part of the marker images is assumed to be known, in order to be used as constraint in the calibration. However, when large numbers of identical markers are used, obtaining a reliable set of stereo pairs in the uncalibrated stereo camera set-up is not a trivial task.

In the present study we developed the Iterative Track Matching (ITM) method to obtain a set of stereo pairs of marker images allowing the calibration of the camera set-up. The ITM method automatically establishes stereo correspondence for large numbers of identical markers (10-100) which are observed in two simultaneously obtained sequences of frames in an uncalibrated X-ray stereo camera set-up. The method assumes that part of the markers is correctly tracked in each of the views, and coarse prior knowledge ($\pm 20^\circ$) of the angle between the two camera axes is used. Ambiguities are solved by identifying corresponding tracks instead of single points. Singular Value Decomposition (SVD) of the set of tracks is used to reduce noise influence and computation time. Next to the set of stereo pairs, the ITM algorithm also renders accurate estimates of the set-up parameters which can be used as initial estimates in the calibration of the set-up [35]. Once the set-up is calibrated, the geometry is accurately known and the set of stereo pairs can be reviewed and extended. Because of its intended use in measuring deformation of the heart, the ITM algorithm is developed for an aperture angle of the object of approximately 0.1 rad.

In computer simulations markers were distributed randomly over the volume of the wall of a moving left ventricle (LV) which was observed in an X-ray stereo camera set-up. For various numbers of attached markers, corresponding sets of tracks in the left and the right view were generated. The ability of the ITM method to identify correct stereo pairs, and the accuracy of the obtained set-up parameter estimates were investigated. Also the performance was evaluated with incomplete sets of tracks which were obtained by random removal of tracks in each of the views. Then only part of the markers is present in both sets, representing the situation of missing tracks due to marker detection or tracking failures. The effect on the performance of the noise level and of the width of the intervals of the set-up parameters was also investigated.

Epipolar constraint for marker tracks

The geometry of the stereo camera set-up is represented by a model which was introduced by us earlier [35]. As shown in Fig. 2, the two cameras are represented by source points S_L and S_R and unit-length camera axes $S_L C_L$ and $S_R C_R$ pointing perpendicularly to the image planes of views L and R . The baseline $S_L S_R$ intersects the image planes in S'_L and S'_R . The original image coordinate systems of the left and right view are indicated by unit vectors \bar{u}_L, \bar{v}_L and \bar{u}_R, \bar{v}_R , respectively. Rotation of these systems in order to align them with the camera axis planes $S_L S'_R C_L$ and $S'_L S_R C_R$, results in coordinate systems \bar{u}_L', \bar{v}_L' and \bar{u}_R', \bar{v}_R' . An image position P_L in view L is represented by image coordinates (u_L, v_L) . Because the camera axis is of unit length, image coordinate u_L is identical with the tangent of the angle between the camera axis $S_L C_L$ and the line from S_L to the image plane position $(u_L, 0)$, and a similar interpretation holds for coordinate v_L . Point P_R in view R is similarly represented by image coordinates (u_R, v_R) . The line BB' indicates the intersection of the image planes. The geometry of the stereo camera set-up is represented by the parameter vector

$$\theta = (\varphi_L, \varphi_R, \nu, \tau, \gamma)^T \quad (1)$$

with the elements of θ defined as indicated in Fig. 2.

When radiopaque markers attached to the LV are observed by two X-ray cameras simultaneously, their movements are recorded in two sequences of T cine or video frames. In order to measure the motion of a set of M markers, in each frame the positions of the markers are detected, and in successive frames corresponding marker images are identified [32]. With a marker, in each view a sequence of image positions $(u, v)_t$, $t=1, \dots, T$, is associated, the track of the marker in the image plane. When, as indicated in Fig. 2, tracks T_s and T_q in views L and R represent a correct stereo pair, ideally, simultaneously measured marker image positions P_L and P_R lie in a common epipolar plane $S_L S_R P_L P_R$. This so-called epipolar constraint can, in terms of the elevation angles β_L and β_R (see Fig. 2), be expressed [35]

$$\beta_L - \beta_R - \gamma = 0 \quad (2)$$

Alternatively, the constraint can be written in the computationally more favourable form of a linear function of the image position coordinates and their cross-products, according to

$$c_0(\theta) + \sum_{i=1}^4 c_i(\theta) u_i + \sum_{i=1}^4 \sum_{j=i+1}^4 c_{ij}(\theta) u_i u_j = 0 \quad (3)$$

$$(u_1, u_2, u_3, u_4) = (u_L, v_L, u_R, v_R)$$

with functions c_i , $i=0,\dots,4$, and c_{ij} , $i=1,\dots,4$, $j=i+1,\dots,4$, defined in Eq. (25), see Appendix A. Because of noise and other imperfections, in general, constraint (3) will not hold exactly, and a small mismatch must be allowed. For a correct stereo pair of tracks, the mismatch should be small in all frames, and an associated mismatch vector \mathbf{m} can be defined

$$\mathbf{m}(\mathbf{u}_1,\dots,\mathbf{u}_4;\theta) = c_0(\theta) + \sum_{i=1}^4 c_i(\theta) \mathbf{u}_i + \sum_{i=1}^4 \sum_{j=i+1}^4 c_{ij}(\theta) \mathbf{u}_i \otimes \mathbf{u}_j \quad (4)$$

$$\mathbf{u}_i = (u_{i,1},\dots,u_{i,T})^T, i=1,\dots,4$$

with $u_{i,t}$, $i=1,\dots,4$, $t=1,\dots,T$, representing the value of the i -th image coordinate in the t -th frame, and $\mathbf{u}_i \otimes \mathbf{u}_j$, a T -vector containing the element to element product of the operand T -vectors. A data reduction with a factor of approximately $3/T$ is obtained by basing the mismatch on only the 3 most significant components of the Singular Value Decomposition (SVD) of the set of all tracks (Appendix B). Then vectors \mathbf{u}_i , $i=1,\dots,4$, in Eq. (4) are replaced by the corresponding approximations

$$\tilde{\mathbf{u}}_i = \sum_{k=1}^3 \lambda_k^{1/2} a_{i,k} \mathbf{b}_k \quad (5)$$

where $\lambda_k^{1/2}$ and \mathbf{b}_k are defined in Appendix B, and $a_{i,k}$ is the element in vector \mathbf{a}_k (Appendix B) which corresponds to the i -th image coordinate of the current track pair. Noting that the vectors \mathbf{b}_k , $k=1,\dots,T$, are orthonormal, an approximation $\tilde{\mathbf{m}}$ of the mismatch can be defined according to

$$\tilde{\mathbf{m}} = \sum_{k=1}^3 (\mathbf{m}(\tilde{\mathbf{u}}_1,\dots,\tilde{\mathbf{u}}_4;\theta) \cdot \mathbf{b}_k) \mathbf{b}_k \quad (6)$$

where \cdot indicates the scalar product of two vectors. The corresponding mismatch measure is defined

$$|\tilde{\mathbf{m}}|^2 = \sum_{k=1}^3 (\mathbf{m}(\tilde{\mathbf{u}}_1,\dots,\tilde{\mathbf{u}}_4;\theta) \cdot \mathbf{b}_k)^2 \quad (7)$$

and for a set of G stereo pairs of tracks the total mismatch is defined

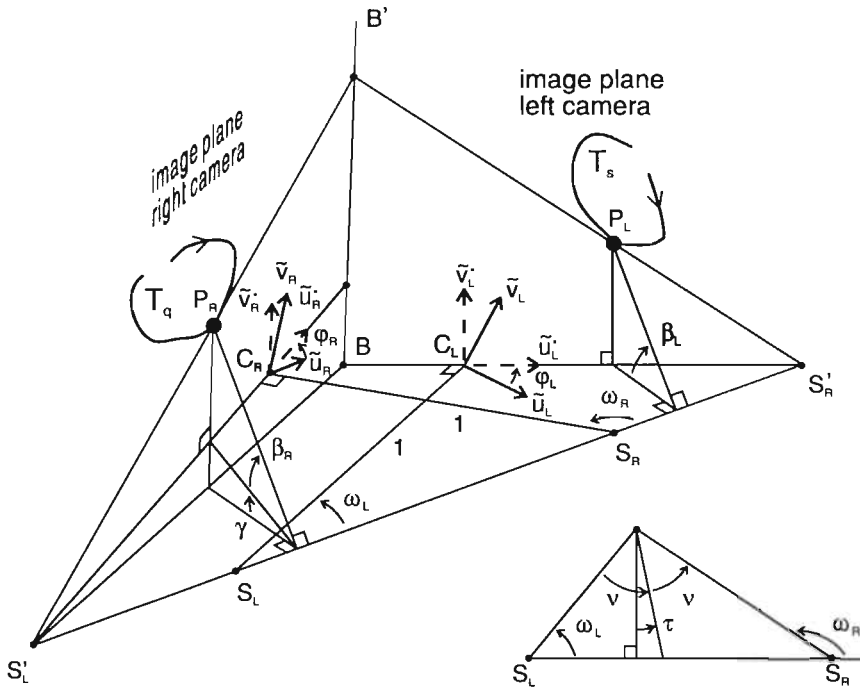


FIGURE 2. Geometric model of a stereo camera set-up. The cameras are represented by source points S_L and S_R and unit-length camera axes $S_L C_L$ and $S_R C_R$ pointing perpendicularly to the image planes of views L and R . The baseline $S_L S_R$ intersects the image planes in S'_L and S'_R . The original image coordinate systems of the left and right view are indicated by unit vectors \tilde{u}_L, \tilde{v}_L and \tilde{u}_R, \tilde{v}_R , respectively. Rotation of these systems in order to align them with the camera axis planes $S_L S'_L C_L$ and $S_L S'_R C_R$, results in coordinate systems $\tilde{u}'_L, \tilde{v}'_L$ and $\tilde{u}'_R, \tilde{v}'_R$. Line BB' indicates the intersection of the image planes. The geometry of the stereo camera set-up is represented by the angular parameters $\phi_L, \phi_R, \omega_L, \omega_R$ and γ , or equivalently, according to the definitions indicated in the triangle right below, by ϕ_L, ϕ_R, v, τ , and γ . Tracks T_s and T_q shown in the left and right view, respectively, form a stereo pair of tracks associated with one marker. Then the simultaneously measured marker images P_L and P_R obey the epipolar constraint: they lie in a plane through source points S_L and S_R , or, equivalently, the corresponding elevation angles β_L and $\beta_R + \gamma$ are equal.

$$m_{tot}^2(\theta) = \sum_{g=1}^G |\tilde{m}_g(\theta)|^2 \quad (8)$$

with \tilde{m}_g , $g=1, \dots, G$, the mismatch vector defined in Eq. (7), for the g -th track pair. The unknown parameter θ can be estimated by minimizing the sum of squares defined in Eq. (8), using a standard nonlinear regression procedure. Substitution of the obtained estimate $\hat{\theta}$ into Eq. (8) results in the final estimate of the mismatch. The corresponding RMS mismatch m per stereo pair of marker images is defined

$$m^2 = \sum_{g=1}^G |\tilde{m}_g(\hat{\theta})|^2 / (G \times T) \quad (9)$$

with G the number of stereo pairs, and T the number of elements of the mismatch vector per stereo pair. For reference purposes it is relevant to know the expected value of the squared mismatch m^2 for correct stereo pairs. In Appendix C it is derived that

$$E(m^2) = (3/T - (n_p / (G \times T)) / 2 \sigma^2 \quad (10)$$

where E is the mathematical expectation operator, n_p the number of fitted parameters, and σ the RMS image position noise.

Iterative Track Matching (ITM) algorithm

A set of stereo pairs of marker tracks is obtained in an iterative procedure (Fig. 3). Initially a starting set of two stereo pairs is obtained, the corresponding set-up geometry is estimated and the set is iteratively extended with stereo pairs which have a small mismatch for the estimated geometry. When extension is stuck while the number of selected stereo pairs is still too small, a new trial is started by selecting a different starting set of two pairs. The algorithm is stopped when a sufficiently large set has been obtained, or the maximum number of trials is reached.

Input, Output. As input the algorithm uses two sets of marker tracks obtained in views L and R (matrix U defined in Eq. (28), Appendix B), the numbers Q and S of tracks in the two sets, the number of markers M , an estimate σ_0 of the RMS noise in the marker image position coordinates, and v_0 , a prior estimate of parameter v . When the algorithm

is successfully completed, the output consists of at least 5 identified stereo pairs of tracks, and the corresponding estimate of the set-up parameter vector θ defined in Eq. (1).

Initialization. For the marker tracks in matrix U , the 3 most significant components of the SVD are calculated (Fig. 3, module 1). When the probability of a track to be missing in one view is the same whether or not the corresponding track in the other view is missing, the expected number of stereo pairs of tracks, n_{LR} , can be estimated by

$$\hat{n}_{LR} = Q \times S / M \quad (11)$$

with Q and S , the numbers of marker tracks in the left and right view, respectively, and M the number of markers.

Iteration. In modules 2 and 3 the initial two pairs of the set of selected stereo track pairs are obtained. First two tracks in view L are selected. Then all possible combinations of two view R tracks are successively matched with the view L couple. Twenty view R couples with minimum mismatch are ordered according to increasing mismatch. The first starting set of two stereo pairs is obtained by combining the current view L couple with the first view R couple in the sequence. If iterative extension of the set of stereo pairs gets stuck too early, a new starting set of two stereo pairs is formed by combining the current view L couple with the next view R couple in the sequence (module 3). When all 20 candidate view R couples have been used without success, a new view L couple and the corresponding sequence of 20 view R couples are obtained (module 2), and iteration proceeds.

In selecting a view L starting couple (module 2), two tracks with different motion patterns are preferred, because they favour reliable matching with view R tracks and accurate estimation of the set-up parameters. The sequence of 20 view R couples with minimal mismatch with the current view L couple is obtained in two stages in order to save computation time. First each possible view R couple is combined with the current view L couple, and for each combination the mismatch for the corresponding set of two stereo pairs is obtained in a linear regression procedure based on Eqs. (26) and (27), Appendix A. Then 50 view R couples with minimum mismatch are selected, and for each selected couple the set-up parameters and the mismatch with the current view L couple are obtained in a nonlinear regression procedure based on Eqs. (7) and (8). The parameter estimates obtained in the linear procedure are used as initial values in the nonlinear procedure.

In module 4, all stereo pairs which can be formed from the tracks which are not present in the current set of selected stereo pairs are considered as candidate pairs. For each candidate pair the RMS mismatch is calculated according to

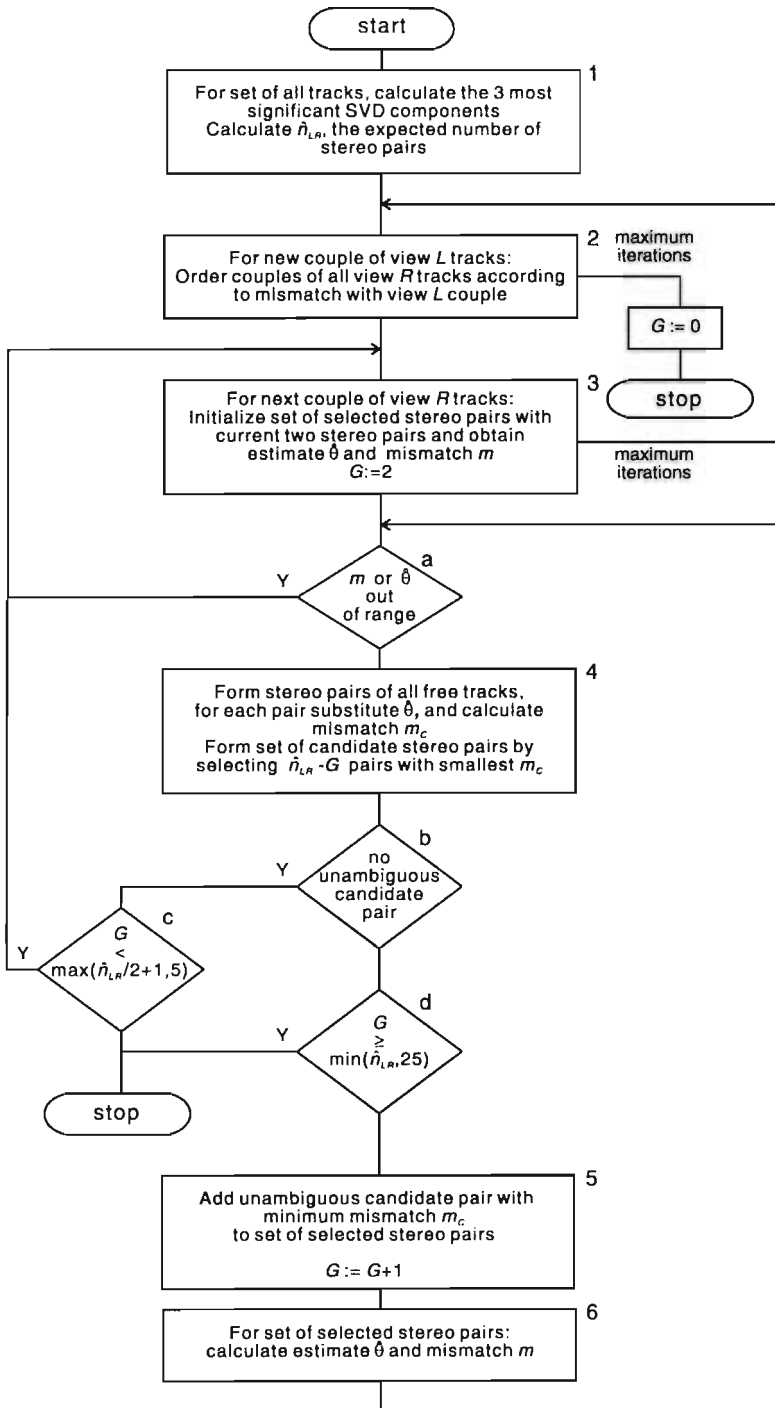
$$m_C = (|\tilde{m}(\hat{\theta})|^2 / T)^{1/2} \quad (12)$$

where $\tilde{m}(\hat{\theta})$ is the measure defined in Eq. (7) with the current estimate $\hat{\theta}$ substituted for θ . So, the obtained mismatch expresses how well an additional stereo pair fits the set-up parameters determined from the current set of selected pairs. A subset of \hat{n}_{LR} - G pairs with the lowest m_c is selected to form the set of candidate stereo pairs. A candidate pair is considered unambiguous when its member tracks do not occur in another pair in the set of candidate pairs. The unambiguous pair with minimum mismatch is added to the set of selected stereo pairs (module 5).

In module 6, similar to module 3, a nonlinear regression procedure is applied to minimize the $G \times 3$ mismatch components defined by Eqs. (7) and (8), with G the number of stereo pairs in the selection. The parameter estimates obtained in the former iteration are used as initial guesses, and new values for parameter estimate $\hat{\theta}$ and mismatch m defined in Eq. (9) are obtained.

When the number of stereo pairs is still small, in particular the estimate of parameter v is inaccurate [35]. Improvement is obtained by initially fixing the parameter to the prior value v_0 , leaving only 4 parameters to be estimated. In order to reduce the influence of possibly introduced systematic errors, fixing v is accompanied by using weights $\lambda_k^{-1/2}$ for the terms in Eq. (7) in the estimation procedure. Then the influence of the 3 SVD components in the estimation becomes more equal, which results in better averaging out of the systematic errors. Therefore, in module 2, and also in module 5 while the number of pairs in the selection is still small (≤ 5), the condition of the parameter estimation is improved by fixing parameter v and using the weights. After 5 pairs have been selected, the constraint on v is relaxed.

FIGURE 3. Flowchart of the Iterative Track Matching (ITM) algorithm. The input of the algorithm consists of two sets of Q and S marker tracks which are simultaneously obtained in two sequences of T frames in views L and R of a stereo camera set-up. The identification of the tracks is unknown, and the goal of the ITM algorithm is to identify correct stereo pairs of tracks. Stereo pairs are selected on the basis of a small epipolar mismatch. Only the motion information in the 3 most significant components of the Singular Value Decomposition (SVD) of the tracks is used in the procedure (module 1). For reference purposes the number of stereo pairs \hat{n}_{LR} is estimated. The set of selected stereo pairs is initialized with two tentative stereo track pairs. The corresponding estimate $\hat{\theta}$ of the stereo camera set-up parameters and the RMS mismatch m are obtained (modules 2,3), and in modules 4,5 the set is extended with an additional track pair which has a small mismatch for $\hat{\theta}$. In module 6 parameters θ and mismatch m are newly estimated for the extended set. This procedure is iterated until a stop criterium is reached (modules b,c,d). When the size G of the final set of selected stereo pairs is considered too small, a new trial search is started in modules 2,3 with a different initial set of two track pairs. The algorithm is stopped with a positive result when G is considered large enough, or with an empty set when the maximum number of trial starting couples in module 2 is reached.



The procedure in modules 4, 5, and 6 is iterated until a stop criterium in modules a-d is reached. The allowed range in module a is defined

$$m < 3\sqrt{E(m^2)} \quad |\hat{v} - v_0| < 20^\circ \quad |\hat{\tau}| < 20^\circ \quad |\hat{\gamma}| < 6^\circ \quad (13)$$

where $E(m^2)$, the expected mismatch for correct stereo pairs, is calculated according to Eq. (10) with σ_0 substituted for σ , and \hat{v} , $\hat{\tau}$, and $\hat{\gamma}$ representing the current estimates of the corresponding parameters. When condition (13) is not obeyed, a new selection is started in module 3. When no unambiguous pairs are present in the set of candidate pairs (module b), and the number G of selected pairs is still considered too small (module c), the selection is also restarted in module 3.

Termination. The algorithm is stopped with a positive result, when the set of selected stereo pairs is considered sufficiently large (modules c,d). The algorithm has failed when the maximum number of trials to obtain an appropriate set of two starting pairs is exceeded (maximum iterations in module 2). When setting the maximum number of trials, it was taken into account that the probability of failing to find a match increases when more tracks are missing. The maximum is defined such that the probability of algorithm failure due to repeatedly selecting a view L couple with no match in view R is less than 0.005.

Implementation. The ITM algorithm was implemented in *FORTRAN* on a VAX8650 computer. For the linear and nonlinear least squares optimization use was made of the NAG Fortran Library routines F01AXF and F04ANF, and E04GDF, respectively. The run-time of the procedure ranged from 20 seconds to 5 minutes for 24 markers and a noise level of 0.5 pixel, and 48 markers and 1 pixel noise, respectively.

Performance evaluation

Computer generation of marker tracks

For the standard X-ray stereo camera set-up used in the evaluation it holds

$$(\varphi_L, \varphi_R, \nu, \tau, \gamma) = (0^\circ, 0^\circ, 45^\circ, 0^\circ, 0^\circ) \quad (14)$$

so, camera axes were perpendicular and intersecting, and object distances were equal. The distance between the camera source points was set at 10^3 mm, corresponding to an object distance of 710 mm. The LV with a diameter of 60-80 mm was positioned with its centre at the intersection of the camera axes, and the long axis randomly oriented in the camera set-up. A number of markers was positioned randomly in the LV wall. Details about the ventricular geometry, the positioning of the markers, and the simulated motion have been presented elsewhere [32]. The motion of the markers in each view

was imaged on video (256×256 pixel, 50 frames), with 1 pixel corresponding to 0.5 mm of the LV. The position noise was set at an RMS value of 0.5 pixel, a value which was motivated previously [32]. The number of markers was equal to 24.

In practice a stereo camera set-up only approximately corresponds to the standard geometry defined in Eq. (14). The corresponding deviations were introduced in repeated simulations by setting the parameters to values obtained by random sampling in the intervals

$$\nu, \tau \in (-5^\circ, 5^\circ) \quad \gamma \in (-1^\circ, 1^\circ) \quad \varphi_L, \varphi_R \in (-45^\circ, 45^\circ) \quad (15)$$

So, the intervals defined in Eq. (15) represent the prior uncertainty of the parameter values. For a set of moving markers observed in a sequence of 50 frames, the exact marker image coordinates were obtained by calculating the projections according to the known viewing geometry. Data generation was completed by adding gaussian noise to the obtained exact marker image coordinates.

Evaluation of the ITM method

In 100 repeated simulations marker tracks were generated as described above. In each simulation new random values were generated for the orientation of the LV, the positioning of the markers in the LV wall, and the set-up parameters. Furthermore, newly generated random noise was added to the exact marker image positions. The prior estimates of the RMS noise level and the set-up parameter ν , the input parameters σ_0 and ν_0 defined in the Section 'ITM algorithm', were in the evaluation set equal to σ , the RMS level of the generated noise, and to 45° , respectively.

The results of the ITM method as obtained in 100 repeated simulations were summarized by counting the number of simulations in which the algorithm failed. Furthermore, for the simulations with successful completion of the algorithm the average numbers of correct and false pairs in the final selection of stereo track pairs, the average mismatch, and the RMS error of the parameter estimates were calculated.

Besides for the standard values presented above, the performance was investigated for varying numbers of markers, levels of the image position noise, and intervals for the parameters which were twice as wide (Table I). Also, the performance was evaluated for complete and incomplete sets of tracks. The incomplete sets were obtained by random removal of a fraction f_m of the generated tracks in each view, resulting in

$$n_L = n_R = M \times (1 - f_m) \quad n_{LR} = M \times (1 - f_m)^2 \quad (16)$$

with n_L and n_R the number of remaining tracks in views L and R , respectively, and n_{LR}

Table I. Settings of the computer simulations in the evaluation of the ITM method.

Intervals of parameters $\varphi_L, \varphi_R, v, \tau, \gamma$	$(\pm 45^\circ, \pm 45^\circ, \pm 5^\circ, \pm 5^\circ, \pm 1^\circ),$ $(\pm 45^\circ, \pm 45^\circ, \pm 10^\circ, \pm 10^\circ, \pm 2^\circ)$
Number of markers (M)	12, 24, 48, 72, 96
RMS noise level (σ)	0.25, 0.5, 1.0 pixel
Fraction tracks missing per view (f_m)	0%, 30%, 50%

the number of stereo pairs. Hence, per view the input to the ITM algorithm contains $n_L - n_{L,R}$ tracks which have no match in the other view.

Results

Tables II and III show the results of the ITM method when applied to computer generated sets of tracks corresponding to 12, 24, 48, 72 and 96 markers attached to a moving LV. Each line of the table represents the results obtained in 100 repeated simulations. Results are shown for the wider parameter intervals according to the second line of Table I, an RMS image position noise level of 0.25, 0.5 and 1.0 pixel, and per view a fraction of 0%, 30% and 50% of the generated tracks considered as missing. For 12 markers the results for 50% missing tracks could not be shown, because the number of 3 remaining stereo pairs is less than the required minimum of 5 stereo pairs of tracks in the ITM method. For numbers of markers up to 48 the results for a RMS noise level of 0.25 pixel were similar (within a few percent) to those for 0.5 pixel, and therefore are not shown. For numbers of markers of 72 and 96 an RMS noise level of 1 pixel was beyond the performance limits of the procedure: the simulations were stopped because of the required excessive amounts of run-time (more than 30 minutes per simulation). The results of simulations with the smaller intervals of the parameters (first line of Table I) were only slightly better and therefore are not shown.

For simulations with 24 markers (Table II, middle panel), a noise level of 0.5 pixel (column 1), and no missing tracks (column 2), the ITM method performed well. In the 100 repeated simulations the algorithm did not fail (column 3) and only rarely a false pair of tracks occurred in the final selection (columns 4,5). The RMS mismatch (column 6), calculated according to Eq. (9), is in agreement with its expected value of approximately 0.2 pixel, as obtained by substituting σ into Eq. (10). Accurate estimates of the parameters were obtained as shown in the last four columns. With 30% of the 24 generated tracks randomly removed in each view, corresponding to only 12 stereo pairs being present in the input set (column 2), the algorithm remained performing well,

although the fraction of false pairs increased to 1% and the accuracy of the parameters decreased. When 50% of the tracks were removed, leaving only 6 stereo pairs in the input, the performance of the algorithm degraded to an inappropriate level, as is shown by the large number of false stereo pairs (column 5) and the poor parameter estimates (columns 7-10). For 24 markers, increasing the noise level to 1 pixel results in a decreased performance of the algorithm as shown by the larger numbers of failures and false pairs, and less accurate parameter estimates. However, with less than 30% of the tracks missing, the algorithm still renders useful results.

In the simulations with 12 markers (upper panel) the algorithm performed well, but with 30% of the tracks missing and a noise level of 1 pixel the performance seriously degraded. With 48 markers (lower panel) and a noise level of 0.5 pixel the algorithm performed well. When the noise level was increased to 1 pixel, the algorithm failed in more than 20% of the simulations. In the remaining cases the algorithm still rendered few false stereo pairs and accurate parameter estimates with 0% or 30% missing tracks, but the algorithm degraded with 50% of the tracks removed in each view. For 72 markers and a noise level of 0.5 pixel, performance also started to degrade with 50% missing tracks (Table III). Then in 20% of the simulations the algorithm failed, but still low numbers of false pairs and accurate parameter estimates resulted in the remaining 80%. Similar results were found for all three levels of missing tracks with 96 markers and a noise level of 0.5 pixel.

Discussion

We developed the Iterative Track Matching (ITM) algorithm to obtain an initial set of stereo pairs of marker images in an uncalibrated stereo camera set-up. Such a set is required when calibration of the stereo camera set-up is based on the stereo correspondence of marker images obtained during measurements [35]. The only prior geometric knowledge used in the procedure is the angle between the two camera axes, which is assumed to be known up to a deviation of $\pm 20^\circ$. It is assumed that the sequences of marker image positions measured in time per view, the so-called marker tracks, are available in both views for at least 5 markers. Ambiguities are solved by identifying corresponding tracks instead of single points (Fig. 1). Singular Value Decomposition (SVD) of the set of tracks is used to reduce noise influence and computation time.

In the model used the geometry of the stereo camera set-up (Fig. 2) can be described by a set of 5 parameters, the angles $(\varphi_l, \varphi_R, \nu, \tau, \gamma)$, and the mismatch of a stereo pair of marker tracks is expressed as a T -vector which is a nonlinear function of the 5 set-up parameters, with T the number of frames. In the ITM algorithm a tentative set of stereo pairs is selected, estimates of the set-up parameters are obtained by minimizing the mismatch, and iteratively the selection is extended by stereo pairs which have a small mismatch for the current parameter estimates. When the extension gets stuck, while the selection is still less than 50% of $\hat{n}_{l,R}$, the expected number of available stereo pairs

TABLE II. The ITM method applied to computer simulations with 12, 24, and 48 markers, attached to a moving LV.

$\sigma^{1)}$	$n_{LR}^{2)}$	$f_f^{3)}$	$CP^{4)}$	$FP^{5)}$	$m^{6)}$	RMS parameter error ⁷⁾			
				$CP+FP$		φ	v	τ	γ
pixel		%		%	pixel	deg	deg	deg	deg
$M=12$									
0.5	12	0	12	0.1	0.2	0.1	1	0.3	0.03
	6	2	5.7	3	0.2	9	4	2	0.3
1.0	12	1	12	0.3	0.3	0.3	3	0.7	0.06
	6	15	5.0	12	0.5	20	7	5	0.7
$M=24$									
0.5	24	0	24	0.1	0.2	0.05	0.6	0.1	0.01
	12	0	12	0.8	0.2	0.1	2	0.4	0.04
	6	6	4.1	30	0.4	30	10	7	1
1.0	24	0	22	0.3	0.3	0.1	2	0.3	0.04
	12	3	11	2	0.4	0.7	6	1	0.2
	6	16	2.4	50	0.8	40	10	11	1
$M=48$									
0.5	48	0	25	0.3	0.2	0.1	2	0.3	0.04
	24	1	23	0.5	0.2	0.1	1	0.1	0.02
	12	9	11	3	0.2	0.4	3	0.6	0.07
1.0	48	27	25	1	0.4	0.2	3	0.6	0.07
	24	23	21	2	0.3	0.4	2	0.5	0.08
	12	21	8	20	0.5	4	10	5	0.4

TABLE III. The ITM method applied to computer simulations with 72, and 96 markers, attached to a moving LV.

$\sigma^{1)}$	$n_{LR}^{2)}$	$f_f^{3)}$	$CP^{4)}$	$FP^{5)}$		RMS parameter error ⁷⁾			
				$CP+FP$		φ	ν	τ	γ
pixel		%		%	pixel	deg	deg	deg	deg
<i>M=72</i>									
0.25	72	0	25	0.3	0.1	0.1	2	0.4	0.03
	36	0	25	0.2	0.1	0.05	1	0.1	0.01
	18	3	18	0.4	0.1	0.05	1	0.1	0.01
0.5	72	5	25	0.9	0.2	0.3	3	0.6	0.07
	36	2	25	0.7	0.2	0.1	1	0.2	0.02
	18	19	16	3	0.2	0.3	3	0.6	0.08
<i>M=96</i>									
0.25	96	0	25	0.5	0.2	0.2	3	0.7	0.08
	48	0	25	0.3	0.1	0.06	0.8	0.2	0.02
	24	0	23	0.4	0.1	0.03	0.5	0.1	0.01
0.5	96	16	25	2	0.3	0.4	4	0.9	0.08
	48	10	25	2	0.2	0.2	2	0.5	0.05
	24	23	21	2	0.2	0.1	2	0.5	0.02

¹⁾ RMS noise level of the marker image position measurement

²⁾ Number of stereo pairs in the input set of marker tracks

³⁾ Fraction of cases of ITM failure in 100 repeated simulations

⁴⁾ Average number of correct stereo pairs of tracks in the selection obtained with the ITM method

⁵⁾ Average percentage of false stereo pairs in the selection

⁶⁾ Average RMS value of the mismatch in the selection of stereo pairs

⁷⁾ RMS value of parameter estimate minus true value in 100 repeated simulations; column indicated φ shows average result of φ_L and φ_R

defined in Eq. (11), the result is considered to be unreliable and a new trial search is started (Fig. 3, modules b, and c). Extension of the set proceeds until the number of selected stereo pairs of tracks is equal to the minimum of \hat{n}_{LR} and 25. Then the set is considered large enough to be used in the calibration of the geometry, e.g. by using the ML method developed by us earlier [35].

A starting set of two stereo track pairs is used because with one track pair the estimated set-up parameters are too inaccurate to reliably select a correct track pair to extend the set. Three starting track pairs would result in still better estimates of the parameters, but also in an increase of the number of trial starting sets by a factor equal to the number of tracks per view.

When selecting a matching view R couple for the starting view L couple of tracks (Fig. 3, modules 2,3), the couple with minimum mismatch is not necessarily the correct one. However, the occurrence of a large number of false matches with a smaller mismatch than the correct pair indicates a lack of selectivity which is likely to result in more frequently selecting a false pair. It was shown empirically that using a maximum number of 20 candidate view R couples is adequate.

Only the 3 major SVD components of the tracks are used when calculating the mismatch and estimating the parameters. The signal to noise ratio of the SVD components of the marker tracks decreases with their rank [34], so the use of higher rank components should be avoided in order to reduce the noise influence. When 4 parameters (parameter v is initially fixed) and the mismatch are estimated for 2 stereo pairs of tracks, at least 3 SVD components are required to provide the minimum of 5 constraints. The use of the 3 SVD components instead of the T original observations reduces the computational effort of the estimation procedure by a factor $3/T$.

The choice of the weights $\lambda_k^{-1/2}$, $k=1,2,3$ which are used for the terms in Eq. (7) when parameter v is fixed in the estimation procedure, is a compromise between using no weights, which is optimal for reducing the noise influence, and weights λ_k^{-1} , $k=1,2,3$, which is optimal to reduce the influence of systematic errors when their relative contribution to each term in Eq. (7) is equal.

A candidate pair is called ambiguous when one of its members also has a small mismatch in combination with another track. For ambiguous candidates the probability that the combination with minimum mismatch is a false one is relatively high. So, by demanding that the candidate pair is unambiguous (Fig. 3 module b), the probability that a false pair enters the selection is reduced.

In the model used (Fig. 2) the cameras are represented as ideal pinhole cameras, and the image distance and the intersection of the camera axis with the image plane are assumed to be known. A real X-ray stereo camera set-up can be represented by this model when the interior calibration [30] of the imaging system is adequate. The major distortions of the imaging system can be corrected by the use of a dewarp grid [3]. The reported residual error of approximately 0.1 pixel is small compared to the RMS noise of 0.5 pixel. Computer simulations showed that for a noise level of 0.5 pixel, the relative error in the mismatch is of the same order as the relative error in the estimate of the image distance, the distance from source point to dewarp grid plane. So, even

with an error of 1%-10% in the image distance, the identification of correct stereo pairs on the basis of a small mismatch is hardly disturbed. It can be derived that an error δ in the position of the intersection of the camera axis with the image plane causes a systematic error of the order of $u^2\delta$ in the angle between two imaging rays, where u represents an image coordinate. With a viewing aperture of 0.1 rad, corresponding to 200 pixels, and assuming $\delta=\pm 0.005$ rad, the error introduced in the image plane is 0.02 pixel, which is small compared to the noise. In conclusion: the systematic errors, which are introduced by inaccurate knowledge of the interior parameters of the imaging systems, in our set-up are negligible compared to the noise level of 0.5 pixel. For substantially lower noise levels, the systematic errors may become dominating, and a more accurate interior calibration of the imaging systems is indicated.

The performance was evaluated in computer simulations of radiopaque markers attached to a moving left ventricle observed in an X-ray stereo camera set-up with approximately perpendicular camera axes. The settings of the simulations were varied as shown in Table I. Two failure measures were used (Tables II,III): f_f , the number of times in 100 repeated simulations that the algorithm failed to obtain a sufficiently large selection of stereo pairs within a maximum number of iterations, and $FP/(CP+FP)$, the average percentage of false pairs in the final selections of stereo pairs obtained in the simulations with a sufficiently large final set.

An increase of the noise level σ resulted in an increase of the mismatch of the correct pairs. This implied more overlap of the mismatch distributions of false and correct pairs, and thus an increased probability of false pairs entering the selection. As a consequence both failure measures increased, see Tables II,III. A similar effect was found when the number of markers M was increased. Then false pairs with a small mismatch are more likely to occur because more tracks have a similar motion pattern.

When the number of missing tracks is increased, false pairs with a small mismatch will also more often enter the selection, because the competing correct pair does not exist. As a consequence, again both failure measures increased (Tables II,III). For 48 and 96 markers and a noise level of 1 and 0.5 pixel, respectively, the effect was not observed for measure f_f because the performance was already severely degraded when no tracks were missing.

Despite the presence of false pairs, the parameters of the set-up in general were accurate. Only for percentages of false pairs larger than 10%, the estimates of in particular the set-up parameters ϕ_L and ϕ_R became inaccurate, with RMS error values of 20° to 40°.

The performance of the ITM algorithm is not sensitive for deviations from the perpendicular stereo set-up geometry defined in Eq. (14), as shown by the similar results which were obtained in simulations with parameters according to the two sets of intervals shown in Table I. Also, performance is not sensitive to the bias which is introduced when initially in the procedure parameter v is fixed to 45° while the true value lies between 35° and 55°.

In recently developed systems for measuring 3-D heart motion by following radiopaque markers [3-5,7] limited numbers of markers (14-22) are used, and stereo

pairs are established in semi-automatic procedures in an already calibrated set-up, or are assumed to be known [31,33]. With the ITM method, stereo pairs are obtained automatically for large numbers of identical markers (10-100) observed in an uncalibrated set-up, even when a substantial part (30%) of the marker tracks is missing in each of the views.

Conclusions

Stereo correspondences of large numbers of moving identical markers (10-100) observed in an uncalibrated stereo camera set-up can automatically be obtained by the use of the Iterative Track Matching (ITM) method. The algorithm uses coarse prior knowledge of the angle between the camera axes ($\pm 20^\circ$) and it is assumed that the sequences of marker image positions measured in time per view, the so-called marker tracks, are available in both views for at least 5 markers. The obtained stereo correspondences can be used as constraints in a calibration of the set-up geometry on the basis of marker image positions obtained during the measurements. Then a separate calibration procedure with a calibration object is no longer required.

In computer simulations with 48 markers observed in 50 frames (256×256 pixel images, root mean square (RMS) position noise of 0.5 pixel), in both views 30% of the generated marker tracks were randomly removed, so only 24 complete stereo pairs remained. With the ITM method on average 23 stereo pairs were obtained, only 0.5% of which were false. Also accurate estimates of the camera set-up parameters were obtained (RMS error 0.02° to 1°) which can be used as initial values in a calibration procedure.

When the number of markers, the noise level or the percentage missing tracks is increased, the performance decreases. With 24 markers, performance remains well up to a noise level of 1 pixel and 30% of the tracks missing, with 48 and 72 markers up to a noise level of 0.5 pixel and 30% missing tracks, and with 96 markers up to a noise level of 0.25 pixel and 50% missing tracks.

Appendix A

Epipolar constraint

Earlier we have shown [35] that the elevation angles β_L and β_R can be specified as a function of the image coordinates (u_L, v_L) , and (u_R, v_R) , and the set-up parameter θ defined in Eq. (1), by the following equations

$$\beta_L = \arctan(v_L^* / r_L) \quad \beta_R = \arctan(v_R^* / r_R) \quad (17)$$

$$r_L = \sin \omega_L - u_L^* \cos \omega_L \quad r_R = \sin \omega_R - u_R^* \cos \omega_R \quad (18)$$

$$\begin{aligned}
 \begin{bmatrix} \dot{u}_L \\ \dot{v}_L \end{bmatrix} &= R_L \begin{bmatrix} u_L \\ v_L \end{bmatrix} & \begin{bmatrix} \dot{u}_R \\ \dot{v}_R \end{bmatrix} &= R_R \begin{bmatrix} u_R \\ v_R \end{bmatrix} \\
 R_L &= \begin{bmatrix} c_L & s_L \\ -s_L & c_L \end{bmatrix} & R_R &= \begin{bmatrix} c_R & s_R \\ -s_R & c_R \end{bmatrix} & \begin{aligned} c_L &= \cos \varphi_L & s_L &= \sin \varphi_L \\ c_R &= \cos \varphi_R & s_R &= \sin \varphi_R \end{aligned}
 \end{aligned} \tag{19}$$

$$\omega_L = \pi/2 - (\nu - \tau) \qquad \omega_R = \pi/2 + (\nu + \tau) \tag{20}$$

The epipolar constraint Eq. (2) can be alternatively written

$$\tan(\beta_L - \beta_R) - \tan \gamma = 0 \tag{21}$$

Application of $\tan(a-b) = \tan(a) - \tan(b) / (1 + \tan(a)\tan(b))$, substitution of Eq. (17), and multiplication with $(1 + \tan(\beta_L)\tan(\beta_R)) \times r_L r_R$ results in

$$\dot{v}_L r_R - \dot{v}_R r_L - (r_L r_R + \dot{v}_L \dot{v}_R) \tan \gamma = 0 \tag{22}$$

Substitution of Eqs. (18), (19), and (20), and for convenience of notation replacing $(\dot{u}_L, \dot{v}_L, \dot{u}_R, \dot{v}_R)$ by $(\dot{u}_1, \dot{u}_2, \dot{u}_3, \dot{u}_4)$, results in

$$c_0(\theta) + \sum_{i=1}^4 c_i(\theta) \dot{u}_i + \sum_{i=1}^4 \sum_{j=i+1}^4 c_{ij}(\theta) \dot{u}_i \dot{u}_j = 0 \tag{23}$$

where the nonlinear functions c_i , $i=0, \dots, 4$, and c_{ij} , $i=1, \dots, 4$, $j=i+1, \dots, 4$, are defined

$$\begin{aligned}
c_0^* &= h \times (\cos 2\tau + \cos 2\nu) \\
c_1^* &= h \times (\sin 2\tau - \sin 2\nu) & c_3^* &= h \times (\sin 2\tau + \sin 2\nu) \\
c_2^* &= 1 - \tan \tau \tan \nu & c_4^* &= -1 - \tan \tau \tan \nu \\
c_{12}^* &= 0 & c_{13}^* &= h \times (-\cos 2\tau + \cos 2\nu) \\
c_{14}^* &= -\tan \tau + \tan \nu & c_{23}^* &= \tan \tau + \tan \nu \\
c_{24}^* &= 2 \times h & c_{34}^* &= 0
\end{aligned} \tag{24}$$

$$\text{with } h = -(\tan \gamma) / (2 \cos \tau \cos \nu)$$

Substitution of Eq. (19) into Eq. (23) results in an identical expression in the measured image position coordinates $(u_1, u_2, u_3, u_4) = (u_L, v_L, u_R, v_R)$, and with functions c_i , $i=0, \dots, 4$, and c_{ij} , $i=1, \dots, 4$, $j=i+1, \dots, 4$, defined

$$\begin{aligned}
c_0 &= c_0^* & \begin{bmatrix} c_1 & c_2 \end{bmatrix} &= \begin{bmatrix} c_1^* & c_2^* \end{bmatrix} R_L & \begin{bmatrix} c_3 & c_4 \end{bmatrix} &= \begin{bmatrix} c_3^* & c_4^* \end{bmatrix} R_R \\
c_{12} &= c_{34} = 0 & \begin{bmatrix} c_{13} & c_{14} & c_{23} & c_{24} \end{bmatrix} &= \begin{bmatrix} c_{13}^* & c_{14}^* & c_{23}^* & c_{24}^* \end{bmatrix} \begin{bmatrix} c_L R_R & s_L R_R \\ -s_L R_R & c_L R_R \end{bmatrix}
\end{aligned} \tag{25}$$

with scalars c_L, s_L, c_R, s_R , and 2×2 matrices R_L and R_R defined in Eq. (19).

By ignoring smaller terms in the epipolar constraint, an approximation is obtained which can be solved by linear regression, thus avoiding the problem of initial estimates. By substituting Eq. (20) into Eq. (18), and the result into Eq. (22), ignoring the smaller terms, and finally substituting Eq. (19), the next linear approximation of the constraint is obtained

$$d_0(\theta) + \sum_{i=1}^4 d_i(\theta) u_i = 0 \tag{26}$$

where the functions d_i , $i=0, \dots, 4$, are defined

$$\begin{aligned}
d_0 &= \tan \gamma (\cos 2\nu + \cos 2\tau) / (2 \cos(\nu + \tau) \cos \varphi_L) \\
d_1 &= \tan \varphi_L \\
d_2 &= -1 \\
d_4 &= (\cos(\nu - \tau) / \cos(\nu + \tau)) \times (\cos \varphi_R / \cos \varphi_L) \\
d_3 &= -d_4 \tan \varphi_R
\end{aligned} \tag{27}$$

Equation (26) represents a linear model with 4 unknown parameters, d_0 , d_1 , d_3 , and d_4 , which can be estimated by applying a standard linear regression procedure. Corresponding estimates for the set-up parameters ϕ_L , ϕ_R , τ , and γ can be derived on the basis of Eq. (27) with prior estimate v_0 substituted for v .

Appendix B

Representing marker tracks by major SVD components

When M markers are observed by two cameras simultaneously in a stereo set-up, ideally, M marker tracks are obtained in each view. In general, however, some marker tracks may be missing due to detection or tracking failures, and unequal numbers of tracks, Q and S , result in views L and R , respectively. So, the resulting set of marker tracks obtained in sequences of T frames can be represented by the $2(Q+S) \times T$ marker position matrix U ,

$$U = \begin{bmatrix} u_{L,1,1} & u_{L,1,2} & \cdots & u_{L,1,T} \\ \vdots & \vdots & & \vdots \\ u_{L,Q,1} & u_{L,Q,2} & \cdots & u_{L,Q,T} \\ v_{L,1,1} & v_{L,1,2} & \cdots & v_{L,1,T} \\ \vdots & \vdots & & \vdots \\ v_{L,Q,1} & v_{L,Q,2} & \cdots & v_{L,Q,T} \\ \\ u_{R,1,1} & u_{R,1,2} & \cdots & u_{R,1,T} \\ \vdots & \vdots & & \vdots \\ u_{R,S,1} & u_{R,S,2} & \cdots & u_{R,S,T} \\ v_{R,1,1} & v_{R,1,2} & \cdots & v_{R,1,T} \\ \vdots & \vdots & & \vdots \\ v_{R,S,1} & v_{R,S,2} & \cdots & v_{R,S,T} \end{bmatrix} \quad (28)$$

where $(u_{L,q,t}, v_{L,q,t})$ is the position of a detected marker in track q in the t -th frame in view L , $1 \leq q \leq Q \leq M$, $1 \leq t \leq T$, and $(u_{R,s,t}, v_{R,s,t})$ similarly defined for track s in view R , $1 \leq s \leq S \leq M$.

The major signal components in matrix U can be obtained by calculating the Singular Value Decomposition [34] (SVD) of U . In the ITM algorithm matrix U is replaced by the approximation \tilde{U} which is obtained by skipping all but the 3 most significant SVD components of U , according to

$$\tilde{U} = \sum_{k=1}^3 \lambda_k^{1/2} \mathbf{a}_k \mathbf{b}_k^T \quad (29)$$

where \mathbf{b}_k^T is the transposed of \mathbf{b}_k , and $\lambda_k^{1/2}$, \mathbf{a}_k , \mathbf{b}_k , $k=1,2,3$, are equal to the (k) -th singular value, and left and right eigenvector [38] of U , respectively. The dimension of the vectors \mathbf{a}_k , and \mathbf{b}_k corresponds to the dimension of columns and rows, respectively, of matrix U . In Eq. (29) the rows of matrix \tilde{U} , which correspond to marker image position coordinates as a function of time, are represented as a weighted sum of 3 component time functions \mathbf{b}_k^T , $k=1,2,3$. The j -th element of vector \mathbf{a}_k represents the relative amplitude of the k -th component function in the j -th row of U , and λ_k represents the total energy contribution of the k -th component to all rows [34].

Appendix C

Expected value of the mismatch

The mismatch T -vector $\tilde{\mathbf{m}}$ for a stereo pair of tracks, as defined in Eq. (6), can be written

$$\tilde{\mathbf{m}} = E(\tilde{\mathbf{m}}) + \varepsilon_{\tilde{\mathbf{m}}} \quad (30)$$

with E the mathematical expectation operator, and T -vector $\varepsilon_{\tilde{\mathbf{m}}}$ representing the noise in $\tilde{\mathbf{m}}$. For correct stereo pairs, the epipolar constraint is obeyed, implying

$$E(\tilde{\mathbf{m}}) = 0 \quad (31)$$

When for an aperture of 0.1 rad, the smaller terms in Eq. (4) are ignored, it holds

$$\varepsilon_{\tilde{\mathbf{m}}} \approx \sum_{k=1}^3 ((\varepsilon_L - \varepsilon_R) \cdot \mathbf{b}_k) \mathbf{b}_k \quad (32)$$

$$\varepsilon_L, \varepsilon_R \text{ i.i.d. } N(\mathbf{0}, \sigma^2 I_T)$$

where T -vectors ε_L , and ε_R , represent the noise in vectors \mathbf{v}_L , and \mathbf{v}_R , which correspond to the image coordinates defined in Eq. (19), Appendix A. Vectors ε_L , and ε_R , are independently and identically distributed (i.i.d.) according to a gaussian distribution with zero mean and a variance matrix as indicated, where I_T is the $T \times T$ unit matrix, and σ is the RMS image position noise. Because of Eq. (31) and the orthonormality of the vectors \mathbf{b}_k , $k=1,2,3$, it holds

$$E(|\tilde{\mathbf{m}}|^2) = E\left(\sum_{k=1}^3 |(\varepsilon_L - \varepsilon_R) \cdot \mathbf{b}_k|^2\right) = 3 \times 2 \sigma^2 \quad (33)$$

When calculating the mismatch in G stereo pairs with parameter estimate $\hat{\theta}$ obtained by minimizing the sum of squares defined in Eq. (8), for the expected value it holds

$$E\left(\sum_{g=1}^G |\hat{m}_g(\hat{\theta})|^2\right) = (G \times 3 - n_p) 2\sigma^2 \quad (34)$$

where n_p is the number of fitted parameters. Accordingly, for the expected value of the mismatch per stereo pair of marker images, m^2 defined in Eq. (9), it holds

$$E(m^2) = ((G \times 3 - n_p) / (G \times T)) 2\sigma^2 \quad (35)$$

References

- [1] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers," *Circulation*, **52**, 859-867 (1975).
- [2] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard," *Circulation*, **61**, 966-972, (1980).
- [3] J.B. Garrison, W.L. Ebert, R.E. Jenkins, S.M. Yionoulis, H. Malcom, G.A. Heyler, A.A. Shoukas, W.L. Maughan, and K. Sagawa, "Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms," *Comput. Biomed. Res.* **15**, 76-96 (1982).
- [4] W.C. Hunter, and E.A. Zerhouni, "Imaging distinct points in left ventricular myocardium to study regional wall deformation," In: J.H. Anderson (ed.), *Innovations in diagnostic radiology*, pp. 169-190 (Springer, London, 1989).
- [5] M.A. Niczyporuk and D.C. Miller, "Automatic tracking and digitization of multiple radiopaque myocardial markers," *Comput. Biomed. Res.* **24**, 129-142 (1991).
- [6] Y. Kong, J.J. Morris and H.D. McIntosh, "Assessment of regional myocardial performance from biplane coronary angiograms," *Am. J. Cardiol.* **27**, 529-537 (1971).
- [7] S.A. MacKay, M.L. Potel and J.M. Rubin, "Graphics methods for tracking three-dimensional heart wall motion," *Comput. Biomed. Res.* **15**, 455-473 (1982).
- [8] M.J. Potel, J.M. Rubin, S.A. MacKay, A.M. Aisen, J. Al-Sadir and R.E. Sayre, "Methods for evaluating cardiac wall motion in three dimensions using bifurcation points of the coronary arterial tree," *Invest. Radiol.* **18**, 47-57 (1983).
- [9] A.A. Young, P.J. Hunter and B.H. Smaill, "Estimation of epicardial strain using the motions of coronary bifurcations in biplane cinéangiography," *IEEE Trans. Biomed. Eng.* **39**, 526-531 (1992).

- [10] P.A. Chevalier, J.R. Rodarte, and L.D. Harris, "Regional lung expansion at total lung capacity in intact vs. excised canine lungs," *J. Appl. Physiol.: Respirat. Environ. Exercise. Physiol.* **45**, 363-369 (1978).
- [11] J.H. Wei, E.A. Hoffman, E.L. Ritman, and E.H. Wood, "Cardiogenic motion of the right lung parenchyma in anesthetized intact dogs," *J. Appl. Physiol.* **58**, 384-391 (1985).
- [12] B.A. Simon and W. Mitzner, "Redistribution of pulmonary ventilation during high-frequency," *Fed. Proc.* **46**, 1424 (1987).
- [13] J.L. Péan, C.J. Chuong, M. Ramanathan and R.L. Johnson, "Regional deformation of the canine diaphragm," *J. Appl. Physiol.* **71**, 1581-1588 (1991).
- [14] S.S. Kramer, J.H. Anderson, J.D. Strandberg and M.W. Donner, "A permanent radiopaque marker technique for the study of pharyngeal swallowing in dogs," *Dysphagia*, **1**, 163-167 (1987).
- [15] K.M. Hiiemae, and A.W. Crompton, "Mastication, food transport, and swallowing". In: M. Hildebrand, D.M. Bramble, K.F. Liem, D.B. Wake (eds.), *Functional vertebrate morphology*, pp. 262-290 (Harvard Univ. Press, Cambridge, MA, 1985).
- [16] J.B. Palmer, E. Tanaka, and A.A. Siebens, "Motions of the posterior pharyngeal wall in swallowing," *Laryngoscope*, **98**, 414-417 (1988).
- [17] N.A. Borghese, and G. Ferrigno, "An Algorithm for 3-D Automatic Movement Detection by Means of Standard TV Cameras," *IEEE Trans. Biomed. Eng.* **37**, 1221-1225 (1990).
- [18] G.W.M.P. Peters, *Tools for the measurement of stress and strain fields in soft tissue*, (Ph.D thesis, University of Limburg, Maastricht, The Netherlands, 1987).
- [19] J. Downs, H.R. Halperin, J. Humphrey, and F. Yin, "An Improved Video-Based Computer Tracking System for Soft Biomaterials Testing," *IEEE Trans. Biomed. Eng.* **37**, 903-907 (1990).
- [20] Y. Yakimovsky, and R. Cunningham, "A system for extracting three-dimensional measurements from a stereo pair of TV cameras," *Comp. Graph. Im. Proc.* **7**, 195-210 (1978).
- [21] J.K. Aggarwal, and N. Nandhakumar, "On the computation from sequences of images - A review," *Proc. IEEE*, **76**, 917-935 (1988).
- [22] M. Jenkin, and J.K. Tsotsos, "Applying temporal constraints to the dynamic stereo problem," *Comp. Vis. Graph. Im. Proc.* **33**, 16-32 (1986).
- [23] S.T. Barnard, and M.A. Fischler, "Computational stereo," *Computing Surveys*, **14**, 553-572 (1982).
- [24] U.R. Dhond, and J.K. Aggarwal, "Structure from stereo - A review," *IEEE Trans. Syst. Man Cybern.* **19**, 1489-1510 (1989).
- [25] Y. Shirai, *Three-dimensional computer vision* (Springer, New York, 1987).
- [26] D. Marr, and T. Poggio, "Cooperative computation of stereo disparity," *Science*, **194**, 283-287 (1976).
- [27] S.T. Barnard, and W.B. Thompson, "Disparity analysis of images," *IEEE Trans. Patt. An. Mach. Intell.* **2**, 333-340 (1980).
- [28] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby, "PMF: A stereo correspondence algorithm using a disparity gradient limit," *Perception*, **14**, 449-470 (1985).
- [29] S.T. Barnard, "Stochastic stereo matching over scale," *Int. J. Comp. Vis.* **3**, 17-32 (1989).
- [30] B.P.K. Horn, *Robot Vision*, (McGraw-Hill, New York, 1986).
- [31] C.E. Metz and L.E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory," *Med. Phys.* **16**, 45-51 (1989).

- [32] A.M.M. Muijtjens, J.M.A. Roos, T. Arts, A. Hasman, and R.S. Reneman, "Tracking markers with missing data by iterative lower rank approximation," (Chpt. 4 of this thesis).
- [33] L.E. Fencil, and C.E. Metz, "Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation", *Med. Phys.* **17**, 951-961 (1990).
- [34] A.M.M. Muijtjens, J.M.A. Roos, T.T. Prinzen, A. Hasman, R.S. Reneman, T. Arts, "Noise reduction in estimating cardiac deformation from marker tracks," *Am. J. Physiol.* **258**, H599-H605 (1990).
- [35] A.M.M. Muijtjens, J.M.A. Roos, T. Arts, and A. Hasman, "3-D Reconstruction of marker positions from stereo images using ML estimation," (Chpt. 5 of this thesis).
- [36] G.A.F. Seber and C.J. Wild, *Nonlinear Regression*, (Wiley, New York, 1989).
- [37] Numerical Algorithms Group, *NAG Fortran Library Manual*, Mark 13, (NAG, Oxford (UK), 1988).
- [38] G.W. Stewart, *Introduction to Matrix Computations*, (Academic Press, London, 1973).



General Discussion and Conclusions

In order to investigate cardiac mechanics in relation to ischemia, regional motion and deformation of the myocardium must be measured with sufficient resolution in time and space. Presently, in our laboratory in open-chest animal experiments optical markers are attached to the anterior free wall of the left ventricle (LV) and their motion is followed by using a video camera [1-4]. In order to investigate long term effects of gradually developing ischemia, or to detect gradual cardiac dilatation or myocardial growth, however, deformation must be measured in the intact animal in 3-D and over much longer periods than one heart cycle. At present, following the motion of implanted radiopaque markers in an X-ray stereo camera set-up seems to be most appropriate for obtaining accurate 3-D motion data of the heart [5,6] in closed-chest dog experiments. Recently developed systems for measuring 3-D heart motion by following radiopaque markers [6-9] use a limited number of markers (14-22), establish correspondences between marker images in semi-automatic procedures, and obtain geometric calibration of the stereo camera set-up in a separate procedure using a calibration object having markers at known positions. In the current thesis methods are developed to automatically follow the motion of large numbers (10-100) of identical markers and to increase the accuracy of the image position measurement. Furthermore, an improved method to calibrate the geometry of the stereo camera set-up on the basis of position data obtained during the measurements [10-12] is developed, supplemented with a method to automatically obtain the required set of stereo pairs of marker images. In computer simulations the developed methods were extensively evaluated and performance limits were obtained.

General Discussion

When a set of radiopaque markers implanted in the LV is observed in time by two cameras simultaneously in an X-ray stereo camera set-up, two sequences of cine or video frames with marker images are obtained. After detection of the marker images, in each frame a set of 2-D marker image positions is obtained. To reconstruct the 3-D marker trajectories from the sets of marker image positions, geometric information on the stereo camera set-up must be obtained. Marker image positions should be identified with physical markers in time as well as in the two views (Fig. 3, Chpt. 1), and the highest accuracy of position should be obtained.

When observing the motion of M markers in a sequence of T cine or video frames, each physical marker is related to a marker track, being the sequence of 2-D marker image positions measured in time. So, the set of measurements obtained in two views can be represented by a $4M \cdot T$ matrix of 2-D image coordinates. From these data the corresponding $3M \times T$ 3-D marker coordinates are estimated, so 25% of the measured 2-D coordinates are redundant. The redundant information is used to identify stereo pairs of marker images and to estimate the 5 parameters which describe the geometry of the stereo camera set-up.

For accurate measurement of regional differences in motion, large numbers of

myocardial markers ($M > 10$) should be used. When tracking many markers that are part of a continuum, the motion of the markers is coherent and the corresponding motion data contain redundant information. This redundancy comes in addition to the aforementioned, inherent redundancy of stereo observations. In the proposed SVD model, position estimates are obtained by lower rank approximation of the marker image coordinate matrix based on the Singular Value Decomposition (SVD) [13]. Orthogonal decomposition of a measurement matrix by SVD has been used before to obtain data or noise reduction [14-18]. When the information in the $4M \times T$ marker image coordinate matrix can be adequately represented by a lower rank matrix of rank L , where $L < \text{minimum}(4M, T)$, a fraction $(1-L/4M) \times (1-L/T)$ of the image position coordinates is redundant. For cardiac motion data L is of the order of 10, so with $T=50$ and $M > 10$, more than 60% of the $4M \times T$ 2-D coordinates is redundant, which substantially exceeds the 25% inherent redundancy of stereo observations.

In the current thesis the SVD model is used to obtain noise reduction in marker position data, to extrapolate marker tracks or estimate missing marker positions, and to follow markers in 2-D. When the marker position coordinate matrix is complete, the lower rank approximation is obtained by direct application of SVD (Chpt. 2). For a matrix with missing data a similar lower rank approximation matrix is obtained in a modified, iterative procedure (Chpt. 3). Missing data may occur due to detection or tracking failures. At an early stage of the tracking process missing data occur in the matrix, because only part of the marker image positions is connected to form a set of subtracks.

Before applying the SVD model to marker position data, in consecutive frames measured marker image positions should be identified with physical markers, thus forming marker tracks. Therefore, in our approach processing of the marker image position data starts with the formation of marker tracks, separately in each view. In the developed tracking method (Chpt. 4), relations between image positions are established by iteratively using lower rank approximation to extrapolate the obtained set of subtracks. While iteration proceeds subtracks are connected to form longer subtracks. In the procedure two marker image positions or subtracks are only connected when the mismatch of the candidate pair is sufficiently small and competing pairs have a significantly larger mismatch. The criterion threshold values in the procedure were set to values resulting in conservatively connecting image positions to form tracks: only when there is little doubt about the correctness a connection is made, and in ambiguous situations a decision is postponed until more information is obtained. In this way large amounts of correct correspondences can be obtained while the fraction of false relations is kept at an acceptable low level.

Stereo correspondence (Chpt. 6) is obtained by iteratively selecting matching pairs of tracks in the stereo camera set-up of unknown geometry. Only a coarse prior estimate ($\pm 20^\circ$) of the angle between the camera axes has to be provided. When calculating the mismatch of candidate pairs, the geometry of the set-up must be approximately known. Therefore, in the procedure the identification of stereo pairs is combined with approximate calibration of the set-up. The calibration is based on a computationally attractive

formulation of the stereo mismatch. Initially, when the number of constraints is still small, the numerical stability of the parameter estimation procedure is enhanced by using an approximate 4 parameter model for the set-up geometry. The use of pairs of tracks instead of pairs of single points in calculating the mismatch enables reliable identification of matching pairs, and noise influence and computation time can be reduced by using the SVD model. Similar to the marker tracking procedure, a conservative strategy is used when selecting stereo pairs of tracks, resulting in high amounts of correct pairs while the fraction of false pairs is kept low.

On the basis of the obtained set of stereo pairs, accurate geometric calibration of the stereo camera set-up is obtained by maximum likelihood (ML) estimation in the 5 parameter nonlinear geometric model of the set-up. The model (Fig. 2, Chpt. 5) is based on ideal pinhole camera geometry. It is assumed that the internal parameters (the distance between source point and image plane, the point of intersection of camera axis and image plane, and the dewarp correction) of the two viewing systems are known or have been independently estimated in a separate calibration procedure. The 5 parameter nonlinear model is preferred because it is parsimonious and the physically interpretable parameters are appropriate to gain insight into the estimation and reconstruction problem. The required initial estimates of the parameters are provided by the approximate calibration which is obtained while establishing the stereo correspondence in the former processing step (Chpt. 6).

In the current approach, the internal parameters of the viewing systems and the geometric parameters of the stereo camera set-up are independently estimated in separate procedures. Thus, the occurrence of correlated parameter estimates is prevented. By calibrating the geometry of the stereo camera set-up with data obtained during the measurements, a separate calibration procedure with a calibration object is avoided. Furthermore, when the stereo camera set-up is (accidentally) changed, a recalibration is relatively simply obtained by calculating new parameter estimates on the basis of the most recent measurements. The $M \times T$ stereo pairs of measured marker image positions represent a large set of constraints which can be used to improve the accuracy of the parameter estimation, and thus the accuracy of the 3-D reconstruction.

On the basis of the obtained accurate ML estimates of the set-up parameters the set of stereo pairs of tracks can be reviewed and extended by iterating the stereo correspondence procedure. When finally the 3-D marker trajectories are reconstructed, the SVD model can be applied to reduce noise and estimate missing data in the obtained $3M \times T$ marker coordinate matrix.

In the current approach detection of false relations and repair techniques are not provided and the processing steps (Fig. 5, Chpt. 1) are executed sequentially without iteration. A method for error detection could be based on a model similar to the SVD model used in filtering and extrapolating marker tracks, and overall performance might improve when the processing steps are iterated. However, with the currently obtained vast amounts of correct and small amounts of false relations there is little to be gained with these additional requirements, so their development and implementation are hardly indicated.

The major part of the evaluation of the developed methods is based on computer simulations of the motion of radiopaque markers attached to a moving LV. The stereo camera set-up used in the simulations had approximately perpendicular camera axes and an aperture of approximately 0.1 rad. Realistic motion data were generated by using a 13 parameter kinematic model for 3-D motion of the LV [19], with the parameters set to values according to the motion observed in a physiological experiment. The geometry of the simulated set-up was chosen in agreement with the set-up used in physiological experiments. When measuring deformation, the uniform translation, rotation and scaling of the observed object are not of interest. So, the displacement corresponding to these motion components is eliminated from the error displacements when calculating the reconstruction error in the computer simulations. The use of simulated data in the performance evaluation is advantageous because: the exactly known true values or correct relations can be used as a reference, large amounts of repeated trials can be executed to reduce the effects of statistical variation, and the performance can be investigated under various conditions. On the other hand when using computer simulations certain measurement errors, in particular systematic errors, may have been omitted falsely.

The methods presented in the current thesis have been primarily developed to be applied in measuring regional myocardial deformation by following radiopaque markers in an X-ray stereo camera set-up. However, the developed techniques are more generally applicable. Currently, they are used to follow feature points in images of the heart obtained with the MRI tagging technique, and to follow optical markers in measuring the strain distribution of the rat gastrocnemius medialis during active contraction.

Conclusions

1. By using a model which is based on Singular Value Decomposition (SVD), the position noise in a set of marker tracks is substantially reduced and accurate extrapolations are obtained when some of the tracks are incomplete. A prior estimate of the root mean square (RMS) noise level σ is used to set a threshold to the number of SVD components used in the signal estimation. Application of the SVD based method in computer simulations of 50 myocardial markers recorded in 50 video frames (256×256 pixels, 50 s^{-1}) resulted in a reduction of the noise to 0.3σ . When applying the extrapolation method to computer simulated data of 54 marker tracks observed in 39 frames and with 10% of the marker image positions removed at random, the missing coordinates were estimated with an error equal to 1.5σ . When in the same set-up a 10% triangular part of the marker positions in the set of tracks was removed, the extrapolation error was equal to 6σ .

2. By using SVD based extrapolations of partly known marker tracks in an iterative procedure, a large set of radiopaque markers can be reliably and automatically tracked, despite the complications caused by crossing of tracks and missing marker image positions. In computer simulations 75 markers attached to a moving left ventricle (LV) were observed in 60 frames with the noise level σ set to 0.5 pixel. Marker images were considered missing, when they approached another marker image at a closer distance than 4 pixels. From the remaining marker image positions (91%) a fraction of 90% became part of correct tracks with a duration of more than 50 frames. A fraction of 10% could not be followed, and 0.3% was part of a false track. Taking into account the increased loss of marker image positions due to detection and tracking failures when the number of markers is increased, a number of 50 to 75 myocardial markers was found to be optimal in a stereo camera set-up.
3. By calibration based on Maximum Likelihood (ML) estimation in a 5 parameter nonlinear geometric model of the stereo camera set-up, accurate, unbiased, and uncorrelated estimates of the set-up parameters are obtained. In computer simulations with $M=25$ markers positioned at the surface of a sphere of 70 mm diameter, with the noise level σ set to 0.5 pixel, the RMS deformation error of the reconstruction was found to be equal to 1 mm. When using multiple frame observations of myocardial markers the parameter errors and the corresponding reconstruction error decrease with a factor $1/\sqrt{T}$, with T the number of frames. However, for an aperture of 0.1 rad the reconstruction does not improve anymore when increasing the number of stereo pairs $M \times T$ above 300, because the level of the reconstruction error inherent in the noise of the image positions is reached. The ML method resulted in RMS reconstruction errors which ranged from 50% to 80% of the levels obtained by using the 8 parameter linear model for calibration of the stereo camera set-up [12].
4. Stereo correspondence of large numbers of moving identical markers (10-100) observed in an uncalibrated stereo camera set-up can be automatically obtained by using the developed Iterative Track Matching (ITM) method. The algorithm uses coarse prior knowledge of the angle between the camera axes ($\pm 20^\circ$) and marker tracks must be available in both views for at least 5 markers. In computer simulations with 48 myocardial markers observed in 50 frames and a noise level σ of 0.5 pixel, in both views 30% of the generated marker tracks were randomly removed. So only 24 complete stereo track pairs remained. With the ITM method on average 23 stereo pairs were obtained, only 0.5% of which were false. Also accurate estimates of the parameters of the stereo camera set-up were obtained (RMS error 0.02° to 1°) which can be used as initial values in calibrating the geometry of the set-up.

References

- [1] T.T. Prinzen, T. Arts, F.W. Prinzen and R.S. Reneman, "Mapping of epicardial deformation using a video processing technique," *J. Biomech.* **19**, 263-273 (1986).
- [2] F.W. Prinzen, C.H. Augustijn, T. Arts, M.A. Alessie and R.S. Reneman, "Redistribution of myocardial fiber strain and blood flow by asynchronous activation," *Am. J. Physiol.* **259**, H300-H308 (1990).
- [3] C.H. Augustijn, T. Arts, F.W. Prinzen and R.S. Reneman, "Mapping the sequence of contraction of the canine left ventricle," *Europ. J. Physiol. (Pflügers Arch.)*, **419**, 529-533 (1991).
- [4] F.W. Prinzen, C.H. Augustijn, M.A. Alessie, T. Arts, T. Delhaas and R.S. Reneman, "The time sequence of electrical and mechanical activation during spontaneous beating and ectopic stimulation," *Eur. Heart J.* **13**, 535-543 (1992).
- [5] N.B. Ingels, G.T. Daughters, E.B. Stinson and E.L. Alderman, "Evaluation of methods for quantifying left ventricular segmental wall motion in man using myocardial markers as a standard," *Circulation*, **61**, 966-972, (1980).
- [6] W.C. Hunter, and E.A. Zerhouni, "Imaging distinct points in left ventricular myocardium to study regional wall deformation," In: J.H. Anderson (ed.), *Innovations in diagnostic radiology*, pp. 169-190 (Springer, London, 1989).
- [7] J.B. Garrison, W.L. Ebert, R.E. Jenkins, S.M. Yionoulis, H. Malcom, G.A. Heyler, A.A. Shoukas, W.L. Maughan, and K. Sagawa, "Measurement of three-dimensional positions and motions of large numbers of spherical radiopaque markers from biplane cineradiograms," *Comput. Biomed. Res.* **15**, 76-96 (1982).
- [8] S.A. MacKay, M.L. Potel and J.M. Rubin, "Graphics methods for tracking three-dimensional heart wall motion," *Comput. Biomed. Res.* **15**, 455-473 (1982).
- [9] M.A. Niczyporuk and D.C. Miller, "Automatic tracking and digitization of multiple radiopaque myocardial markers," *Comput. Biomed. Res.* **24**, 129-142 (1991).
- [10] H.C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, **293**, 133-135 (1981).
- [11] C.E. Metz and L.E. Fencil, "Determination of three-dimensional structure in biplane radiography without prior knowledge of the relationship between the two views: Theory," *Med. Phys.* **16**, 45-51 (1989).
- [12] L.E. Fencil, and C.E. Metz, "Propagation and reduction of error in three-dimensional structure determined from biplane views of unknown orientation," *Med. Phys.* **17**, 951-961 (1990).
- [13] G.W. Stewart *Introduction to Matrix Computations*, (Academic, London, 1973).
- [14] H.C. Andrews and C.L. Patterson, "Singular value decompositions and digital image processing," *IEEE Trans. Acoust. Sp. Sign. Proc.* **24**, 26-53 (1976).
- [15] H.C. Andrews and C.L. Patterson, "Singular value decomposition (SVD) image coding," *IEEE Trans. Commun.* **24**, 425-432 (1976).
- [16] A.A. Damen and J. van der Kam, "The use of the singular value decomposition in electrocardiography," *Med. & Biol. Eng. & Comput.* **20**, 473-482 (1982).
- [17] D.W. Tufts, R. Kumaresan and I. Kirsteins, "Data adaptive signal estimation by singular value decomposition of a data matrix," *Proc. IEEE*, **70**, 684-685 (1982).
- [18] G.J.H. Uijen, A. Heringa and A. van Oosterom, "Data reduction of body surface potential maps by means of orthogonal expansions," *IEEE Trans. Biomed. Eng.* **31**, 706-714 (1984).

- [19] T. Arts, W.C. Hunter, A. Douglas, A.M.M. Muijtjens, and R.S. Reneman, "Description of the deformation of the left ventricle by a kinematic model," *J. Biomech.* **25**, 1119-1127 (1992).

Summary

The heart is a hollow, thick-walled muscle, which pumps the blood through the vascular system. Blood and oxygen are supplied to the heart by the coronary arteries. Stenotic atherosclerosis may result in inadequacy of the supply because of (partial) occlusion of the arteries. Regional myocardial ischemia is induced and myocardial function is restricted. The reduced functioning in the ischemic region is reflected by reduced segment shortening or even segment lengthening.

In physiological experiments cardiac mechanics are investigated in relation to occlusion of the coronary arteries. Regional myocardial deformation can accurately be measured in 3-D by following implanted radiopaque markers simultaneously by two 'cameras' in an X-ray stereo camera set-up. Thus two sequences of frames with marker images are obtained. The position of a marker image in a frame provides information about the direction of the markers spatial location as seen from the viewpoint of the camera. A simultaneously obtained image of the same marker with a second camera renders a second direction. Then the 3-D position of the marker is found at the intersection of the two direction rays.

When following a large set of identically shaped markers a major problem is the correct identification of marker images corresponding to the same marker in different frames. Similar identification problems occur when identifying corresponding images in consecutive frames obtained with one camera, and when identifying stereo pairs of marker images in frames which are simultaneously obtained by two cameras. Correct identification is hampered by missing marker images due to detection failure. Besides, detection and identification problems arise when marker images approach closely, or overlap in the case of crossing marker trajectories.

In order to reconstruct the position of a marker from a given stereo pair of marker images, the geometry of the stereo camera set-up must be known. Reconstructed marker trajectories should be as complete and accurate as possible. So, the noise in the position measurement must be reduced as much as possible, and accurate estimates must be provided for the missing positions.

Recently developed methods to follow radiopaque markers in an X-ray stereo camera set-up show some drawbacks:

- a limited number of markers is used (14-22)
- the redundancy in the position data is only incidentally used

- the geometry of the stereo camera set-up is obtained in a separate calibration procedure using recordings of a calibration object
- stereo pairs of marker images are obtained with operator intervention in semi-automatic procedures

In the current thesis methods are developed to accurately and automatically reconstruct the 3-D trajectories of large amounts of markers (10-100). In computer simulations the developed methods were extensively tested and evaluated, and their performance limits were obtained.

An important aspect of our approach is the more integral use of the redundancy in the position data of a set of markers. The movements of markers attached to the heart are to a certain extent mutually dependent and can be represented as combinations of a limited amount of components. The motion of the heart abruptly changes in certain episodes of the heart cycle. With regards to this behaviour, a decomposition is used which distinguishes the coherent part in the marker movements from the noisy, incoherent part, without imposing smoothness restrictions to the components. By application of the Singular Value Decomposition (SVD) the desired decomposition can be obtained.

The sequence of image positions of a marker in consecutive frames is called the track of the marker. For a given set of tracks the SVD components can be obtained by applying standard mathematical techniques. Based on prior knowledge of the noise level in the position measurement, part of the components is removed and the tracks are recalculated on the basis of the contributions of the remaining components. In computer simulations of 50 markers observed in a sequence of 50 video frames (256×256 pixels, 50 s^{-1}) application of the method resulted in a reduction of the noise to 30% of its original RMS (root mean square) value of 0.1 pixel (*Chapter 2*).

Generally, in part of the frames some of the marker images may be missing and as a consequence the corresponding tracks are only partly known. For a set of tracks which contains some incomplete tracks, an iterative version of the SVD method is developed. With this method accurate extrapolations of incomplete tracks can be obtained over wide gaps in time. In computer simulations with the tracks of 54 markers observed in 39 frames, in about 50 % of the tracks a consecutive part of 3% up to 44% of the positions was removed. The SVD method rendered extrapolations with an accuracy of 0.5 pixel (RMS), which is 5 times as high as the noise level, but still small when compared to the total excursion in a track of 20 pixels (*Chapter 3*).

Based on this extrapolation technique a method is developed to reconstruct tracks from the set of unidentified image positions obtained with a camera in a sequence of frames. The procedure starts with connecting image positions in consecutive frames by the use of conventional tracking methods. When marker images approach too closely, reliable decisions about the correct extension of a track cannot be obtained with these methods,

and tracking is interrupted. The result of this procedure is a set of track parts (subtracks) which still have to be identified. By using the extrapolation method described in Chpt. 3 matching subtracks are searched and connected to form larger compound tracks. In computer simulations of markers attached to a naturally moving heart, the motion of 75 markers was observed in 60 frames. In these simulations the noise level was set to 0.5 pixel (RMS). Marker images were assumed to be non-detectable when their mutual distance was less than 4 pixels, which resulted in a loss of 9% of the image position data. Despite this complication, 90% of the markers was correctly followed. For the remaining 10% the subtracks could not reliably be connected, and a fraction of 0.3% of the compound tracks contained a false connection of subtracks associated with different markers (*Chapter 4*).

The geometry of a stereo camera set-up can also be calibrated on the basis of a set of stereo pairs of marker images obtained during the measurements. In our approach existing methods are improved. The geometry is represented as a (nonlinear) model with the minimum of 5 parameters, and Maximum Likelihood (ML) parameter estimates were derived. The improved estimation of the geometry results in smaller errors when reconstructing the 3-D marker positions. In computer simulations 25 markers were distributed over the surface of a sphere (diameter 70 mm) which was located at about 700 mm from the X-ray source. With an image position noise of 0.5 pixel (RMS), the resulting reconstruction error was 1 mm. By using the repeated observations of markers in a sequence of T frames, the reconstruction error can be reduced with a factor \sqrt{T} . By using the ML method the reconstruction error was reduced to 50% to 80% of the errors obtained with the existing methods (*Chapter 5*).

Analysis of the parameter estimation in the geometric 5 parameter model indicates that accuracy problems in the calibration were mainly due to one poorly determined parameter, namely the angle between the camera axes. This knowledge was used when developing a method to automatically obtain the set of stereo pairs of marker images which is required for the calibration. Ambiguities were avoided by searching for matching pairs of tracks instead of pairs of single positions. Reduction of the noise influence as well as the computational effort was obtained by replacing the tracks by an approximation based on the 3 most significant SVD components. The developed algorithm uses coarse knowledge of the angle between the camera axes ($\pm 20^\circ$) and it is required that for at least 5 markers the tracks for both cameras are available. In computer simulations of a naturally moving heart, 48 attached markers were observed in 50 frames, with an image position noise of 0.5 pixel (RMS). Problems with detection or tracking of markers were simulated by removing 30% of the generated tracks for each of the cameras. Thus only 24 markers were observed by both cameras. Besides, each camera observed an additional number of 10 markers which were not observed by the other camera. On average the procedure identified 23 pairs of tracks of which 0.5% were

incorrect pairs. Besides, for the 5 geometric parameters accurate estimates were obtained which can be used as initial estimates in the calibration method described in Chpt. 5 (*Chapter 6*).

With the methods developed in the current thesis, the 3-D motion of large sets of markers attached to the heart can accurately and automatically be reconstructed from the sets of unidentified marker images which are simultaneously obtained with two cameras in an X-ray stereo set-up. The geometry of the stereo camera set-up is calibrated by using marker images obtained during the measurements. As a result, images of a calibration object are no longer required and it is allowed to change the stereo camera set-up between measurement sessions. The stereo pairs of marker images required for the calibration can automatically be obtained in an arbitrary camera geometry of which only the angle between the camera axes is roughly known ($\pm 20^\circ$).

The methods presented in the current thesis were primarily developed to facilitate the measurement of regional cardiac deformation by following radiopaque markers in an X-ray stereo camera set-up. However, the obtained methods are more generally applicable and are currently used in the analysis of images of the human heart obtained by the Magnetic Resonance Imaging (MRI) tagging technique, and in following optical markers in measuring the strain distribution in skeletal muscles.

Samenvatting

Het hart is een holle dikwandige spier die door regelmatig samen te trekken het bloed via de bloedvaten door het lichaam pompt. Via de kransslagaders wordt het hart zelf van bloed en zuurstof voorzien. Door atherosclerose kunnen in deze vaten vernauwingen optreden waardoor een deel van de hartspier slechter doorbloed wordt. Dat spiergedeelte brengt minder samentrekkende kracht op en levert daardoor een geringere bijdrage aan de pompfunctie van het hart. Het functieverlies komt tot uitdrukking in een verminderde samentrekking of zelfs oprekking van het betreffende spiersegment in de fase waarin het hart samentrekt.

In fysiologische experimenten wordt het proces van het dichtgaan van de kransslagaders in relatie tot een plaatselijk verminderde samentrekking van het hart onderzocht. De lokale ruimtelijke beweging en vervorming van het hart kan nauwkeurig gemeten worden door het volgen van geïmplanteerde metalen bolletjes (markeerders) met twee 'camera's' in een stereoscopische röntgenopstelling. Aldus worden twee reeksen beelden verkregen waarin de markeerders als röntgenschaduwen zijn afgebeeld. Aan de hand van de positie van de schaduw in het beeld is bekend in welke richting, gezien vanuit de camera, de markeerder zich in de ruimte bevindt. Een tegelijkertijd verkregen opname met een tweede camera vanuit een andere invalshoek levert een tweede richting op. De ruimtelijke positie van de markeerder komt dan overeen met het punt waar de twee richtingen elkaar snijden.

Wanneer meerdere markeerders, allemaal bolletjes van dezelfde afmetingen, tegelijkertijd worden afgebeeld is het niet eenvoudig om te bepalen welke schaduwen in verschillende beelden van dezelfde markeerder afkomstig zijn. Dat geldt zowel bij het volgen van markeerders in opeenvolgende beelden, als bij het opsporen van stereoscopische schaduwparen in tegelijkertijd met twee camera's verkregen beelden. Het correct identificeren van overeenkomende schaduwen wordt bemoeilijkt door onvolledigheid van de gegevens: markeerders kunnen tijdelijk uit het zicht verdwijnen of hun schaduwen kunnen slecht herkenbaar zijn. Daarnaast kunnen herkenning- en identificatieproblemen ontstaan doordat markeerderschaduwen in het afbeeldingsvlak te dicht bij elkaar komen of (gedeeltelijk) gaan overlappen wanneer afgebeelde markeerdertrajectoriën elkaar kruisen.

Om uitgaande van een bekend stereoscopisch schaduwpaar de ruimtelijke positie van de bijbehorende markeerder te reconstrueren moet de geometrie van de cameraopstelling bekend zijn. Daarnaast is het voor de nauwkeurigheid van de gereconstrueerde markeerdertrajectoriën van belang dat de meetruis zoveel mogelijk onderdrukt wordt en onvolledige positiegegevens aangevuld worden met betrouwbare schattingen.

Recent ontwikkelde methoden voor het volgen van geïmplanteerde markeerders in een stereoscopische röntgenopstelling kennen een aantal beperkingen:

- het gebruikte aantal markeerders is relatief gering (14-22)
- de redundantie in de positiegegevens wordt slechts incidenteel benut
- de geometrie van de cameraopstelling wordt bepaald in een speciale calibratie-procedure waarbij opnamen worden gemaakt van een calibratieobject
- stereoscopische schaduwparen worden verkregen met behulp van interventie van een 'operator' in semi-automatische procedures.

Het promotieonderzoek was gericht op de ontwikkeling van methoden waarmee de ruimtelijke trajectoriën van grote aantallen (10-100) markeerders automatisch en nauwkeurig konden worden gereconstrueerd. In computer simulaties zijn de ontwikkelde methoden uitgebreid getest en geëvalueerd en zijn de grenzen aan hun toepasbaarheid bepaald.

Een belangrijk element in onze aanpak is het beter benutten van de redundantie in positiegegevens van markeerders. De bewegingen van in het hart aangebrachte markeerders zijn in zekere mate afhankelijk van elkaar en kunnen als combinaties van een beperkt aantal componenten gerepresenteerd worden. Voor het hart geldt dat in sommige episodes de beweging abrupt verandert en in verband daarmee is gekozen voor een ontbinding die het coherente deel in de markeerderbewegingen onderscheidt van het incoherente ruisdeel, zonder restricties op te leggen aan de gladheid van de componenten. Door toepassing van Singular Value Decomposition (SVD) wordt een dergelijke decompositie verkregen.

De reeks posities van een markeerder in opeenvolgende beelden wordt een spoor genoemd. Voor een verzameling sporen kunnen de SVD componenten verkregen worden door toepassing van standaard wiskundige technieken. Op grond van kennis van het nivo van de ruis in de positiemeting wordt een deel van de componenten verwijderd en de sporen worden herberekend aan de hand van de gehandhaafde componenten. In computer simulaties van 50 aan het hart gehechte markeerders geobserveerd in 50 videobeelden (256×256 pixels, 50 s^{-1}) leidde toepassing van deze methode tot een reductie van de ruis tot 30% van zijn oorspronkelijke RMS (root mean square) waarde van 0.1 pixel (*Hoofdstuk 2*).

Vanwege het ontbreken van posities in een deel van de reeks beelden kunnen sommige sporen slechts gedeeltelijk bekend zijn. Voor een verzameling sporen die niet allemaal volledig bekend zijn, is een iteratieve variant van de SVD methode ontwikkeld. Met behulp daarvan kunnen voor de onvolledige sporen nauwkeurige extrapolaties verkregen worden over vele beelden heen. In computer experimenten met sporen van 54 aan het hart gehechte markeerders geobserveerd in 39 videobeelden werd bij ongeveer de helft van de sporen een aaneengesloten gedeelte van 3% oplopend tot 44% verwijderd. De met de SVD methode verkregen extrapolaties hadden een nauwkeurigheid van 0.5 pixel (RMS), wat een factor 5 hoger is dan het ruisnivo maar nog steeds klein in

vergelijking met de totale excursie van een marker die ongeveer 20 pixels bedraagt (Hoofdstuk 3).

Op basis van deze extrapolatietechniek worden sporen samengesteld uit de verzameling ongeïdentificeerde schaduwposities in een reeks beelden. De procedure start met het aan elkaar rijgen van de bij eenzelfde markeerder behorende schaduwposities met behulp van conventionele volgtechnieken. Wanneer schaduwen elkaar te dicht naderen kan met deze methoden geen betrouwbare beslissing over de juiste voortzetting genomen worden. Dan wordt een tot dan toe gevolgd spoor beëindigd. Het resultaat van deze verwerking is een verzameling losse spoorstukken waarvan de identiteit nog onbekend is. Op basis van de extrapolatiemethode uit Hfdst. 3 wordt vervolgens gezocht naar goed bij elkaar passende spoorstukken, die gekoppeld worden tot langere, samengestelde sporen. In computer simulaties van markeerders gehecht aan een realistisch bewegend hart werden 75 markeerders geobserveerd in 60 frames. Bij deze simulaties van observaties werd uitgegaan van een ruisnivo van 0.5 pixel (RMS). Markeerder-schaduwen werden beschouwd als niet detecteerbaar indien de onderlinge afstand kleiner was dan 4 pixels, hetgeen leidde tot verlies van 9% van de positiegegevens. Ondanks deze complicatie werd 90% van de markeerders correct gevolgd. Voor de overige 10% konden de spoorstukken niet betrouwbaar gekoppeld worden, en een fractie van 0.3% van de samengestelde sporen bevatte een incorrecte koppeling tussen spoorstukken van verschillende markeerders (Hoofdstuk 4).

De geometrie van de cameraopstelling kan ook gecalibreerd worden op basis van een verzameling stereoscopische schaduwparen die verkregen is tijdens de metingen. Bestaande methoden zijn in onze aanpak verbeterd. Daarbij wordt de geometrie gerepresenteerd als een (niet-lineair) model met het minimale aantal van 5 parameters, waarvoor Maximum Likelihood (ML) schatters zijn afgeleid. De aldus verkregen verbeterde schatting van de geometrie leidt tot kleinere fouten bij de reconstructie van de ruimtelijke positie van de markeerders. In computer simulaties werden 25 markeerders verdeeld over het oppervlak van een bol (diameter 70 mm) die op ongeveer 700 mm van de röntgenbron geplaatst was. Een ruis in de positiemeting van 0.5 pixel (RMS) resulteerde in een reconstructiefout van 1 mm. Door herhaling van de waarnemingen van de markeerders in een reeks van T beelden kan de reconstructiefout met een factor \sqrt{T} verkleind worden. Met de ML methode werd de reconstructiefout verlaagd tot 50% à 80% van de fout zoals deze gevonden is bij toepassing van bestaande methoden (Hoofdstuk 5).

Analyse van het 5 parameter model wijst uit dat problemen bij de calibratie met name veroorzaakt worden door het slecht bepaald zijn van één van de 5 parameters van de camera opstelling, namelijk de hoek tussen de camera-assen. Van die kennis is gebruik gemaakt bij de ontwikkeling van een methode voor het automatisch opsporen van de voor de calibratie noodzakelijke verzameling stereoscopische paren schaduwposities. Door te zoeken naar paren sporen in plaats van paren posities zijn correcte en incorrecte

stereoscopische paren beter van elkaar te onderscheiden. Reductie van zowel de ruisinvloed als de rekentijd wordt bereikt door sporen te vervangen door een benadering gebaseerd op de drie meest significante SVD componenten. Het ontwikkelde algoritme gebruikt ruwe kennis van de hoek die de twee camera-assen met elkaar maken ($\pm 20^\circ$) en voor tenminste 5 markeerders dienen de twee sporen beschikbaar te zijn. In computer simulaties werden 48 aan een bewegend hart gehechte markeerders geobserveerd in 50 beelden bij een ruisnivo van 0.5 pixel. Het ontbreken van sporen werd gesimuleerd door per camera 30% van de aangemaakte sporen te verwijderen. Dit had tot resultaat dat 24 markeerders met beide camera's waargenomen werden. Daarnaast nam elke camera nog 10 markeerders waar die niet door de andere camera gevolgd werden. Gemiddeld werden 23 spoorparen gevonden, waarvan een fractie van 0.5% incorrect was. Voor de 5 geometrische parameters werden nauwkeurige schattingen verkregen die als beginschattingen kunnen worden gebruikt in de in Hfdst. 5 beschreven calibratiemethode (*Hoofdstuk 6*).

Met de in het promotieonderzoek ontwikkelde methoden kunnen de ruimtelijke bewegingen van grote aantallen (10-100) aan het hart gehechte markeerders automatisch en nauwkeurig worden gereconstrueerd, uitgaande van de verzamelingen ongeïdentificeerde schaduwposities in simultaan met twee camera's in een stereoscopische röntgenopstelling verkregen opnamen. De cameraopstelling wordt gecalibreerd op basis van schaduwposities verkregen tijdens de metingen. Als gevolg daarvan zijn geen extra opnamen meer nodig van een calibratieobject en kan de opstelling zonder bezwaar tussen verschillende metingen gewijzigd worden. De voor de calibratie benodigde stereoparen van schaduwposities worden automatisch bepaald. Daarbij wordt alleen gebruik gemaakt van ruwe kennis ($\pm 20^\circ$) van de hoek tussen de camera-assen in een verder onbekende geometrie.

De in dit proefschrift beschreven methoden zijn primair ontwikkeld om gebruikt te worden bij het meten van de lokale vervorming van het hart door het volgen van metalen markeerders in een stereoscopische röntgenopstelling. De methoden zijn echter algemener inzetbaar en worden momenteel gebruikt bij de verwerking van beelden van het menselijk hart verkregen met de Magnetic Resonance Imaging (MRI) tagging techniek en bij het volgen van optische markeerders voor het meten van de rek-distributie in skeletspieren.

Resumé

Umtot 't Mestreechs ziech in 't algemeijn get langer trèk daan 't Ingels of 't Nederlands hōbbe veer dit resumé beperrek tot de probleembesjreijving en de algemeijn kōnkluzie. Detajs euver d'n aonpak en de oontwikkelde methodes kint geer daan nao kääs in eijn vaan die twie aander taole tot uuch numme.

't Hart is 'n hool dikwandige spier die zich regelmaotig sametrèk en daomèt deveur zōrrig tot 't blood via de aojers door 't liechaam weurt gepoomp. Via de kransslaagaojers weurt 't hart zellef veurzeen vaan blood en zuurstof. Door ateroskleroos kinne in die aojers vernawwinge goon optreije mèt es gevolg tot e stökske vaan 't hart slechter doorblooid weurt. Dat spiergedeilte bringk daan minder sametrèkkende krach op boedoor 'n kleinder beijdraag weurt gelieverd aon de poompfunctie vaan 't hart. 't Funksieverlees kump tot oetdrökking in e verminderd sametrèkke of zelfs opgerèk weure vaan dat spiersegmint in de episod boein 't hart ziech sametrèk.

In fysiologische experiminte weurt 't perces van towgoon vaan de kransslaagaojers in relasie tot 'n plaotselik verminderde sametrèkking vaan 't hart oonderzoek. 't Rezjo-naal ruimtelik boezjere en vervōrreme vaan 't hart kint gans perceijs gemete weure door 't vollege vaan geïmplanteerde metaole bōlkes (markeerders) mèt twie 'kamera's' in 'n stereoskopische röntgenopstelling. Zoe kreijgste twie series pleetsjes boein de markeerders es sjaduwe te zien zien. Aon de hand vaan de plaots vaan 'ne sjaduw in de afbeelding kinte bepaole in welleke riehting, beloerd vaanoet de kamera, de markeerder moot zitte. Hōbste daoneve nog 'n tegeliektied gemaakte afbeelding mèt 'nne twiede kamera vaanoet 'nnen aanderen hook, daan lievert dat 'n twiede riehting op. De markeerder moot daan zjus zitte boe die twie riehtinge ziech tōtsje.

Es 'n ganze raffel markeerders, allemaol eve groete bōlkes, tegeliektied weurt aafgebeeld, daan is 't neet gemékelek um oet te make welleke sjaduwe in versjèllende pleetsjes aofkomstig zien vaan dezellefde markeerder. Dat deit zich veur beij 't vollege vaan markeerders in opeinvollegende pleetsjes, mer zjus zoe good beij 't beijein zeumere vaan stereoskopische sjaduwpare in twie tegeliektied mèt twie kamera's gemaakte afbeeldinge. 't Zjus identifisere vaan euvereinkoumende sjaduwe weurt bemeujelik door 't neet kompleet zien vaan de gegeves: markeerders kinne effekes oet 't ziech verdweijne of hunne sjaduw kin slech herkinbaar zien. Daoneve kinne herkinnings- en identifikasieprobleme ouch nog oontstoon doortot markeerdersjaduwe te kort beijein koume of (gedeiljtelik) goon euverlappe es de aafgebeelde markeerder-trazjèkke ziech kruse.

Um oetgaond vaan e bekind stereoskopisch sjaduwpaar de ruimteleke peziezie vaan de bejbehurende markeerder te rekonstruere moot de zjeometrie vaan de kamera-opstèlling bekind zien. Daoneve is 't veur de perceijsheid vaan de gerekonstruweerde markeerdertrazjèkke vaan importantie tot de roesj in de metinge zoeväöl meugelik onderdrök weurt en tot oonvolledige pezieziegegeves zoeväöl meugelik aongevöld weure mèt betrowbaar sjattinge.

Recent oontwikkelde methodes veur 't vollege vaan geïmplanteerde markeeders in 'n stereoskopische röntgenopstèlling höbbe 'n aantal beperkinge:

- 't gebruik aantal markeeders is relatief kleijn (14-22)
- de redundantie in de pezieziegegeves weurt mer incidenteel benut
- de zjeometrie vaan de kamera opstèlling weurt bepaold in 'n spesjaal kalibrasieproceduur boebei opnames weure gemaak vaan 'nne kalibrasieklómel
- stereoskopische sjaduwpare weure hallef-otomaties beijeingezeuk mèt eine achter de knup deeje moot ingriepe es 't misgeit.

't Onderzoek in dees dissertasie waor trop geriech um methodes oet te prakkezere en inein te fistermölle boemèt de trazjèkke vaan groete aantalle (10-100) markeeders otomaties kinne weure gerekonstruweerd. In computer simmelasies zien de technieke oetgebrijd getes en getoek zoetot me wèt wat me demèt veerdig kreijg en ouch wienie de zaak in de sop geit loupe umtot de grenze zien bereik.

Wiegezag veur detajs euver d'n aonpak en de oetwèrreking daovaan in de versjèllende hoofstökke vaan dees dissertasie weurt geer verweze nao eijn vaan de twie aander resums. Veer beslete heij mèt de algemeijn kónkluzie.

Mèt de oontwikkelde methodes is 't inderdaad meugelik um de ruimteleke beweginge vaan groete aantalle (10-100) markeeders otomaties en gans perceijs te rekonstruere, oetgaond vaan de hamfele ongeïdentificeerde sjaduwpeziezies in tegelikertied met twie kameras's verkregen series aofbeeldinge. Daobeij weurt 'n belangrieke rol gespäöld door 't SVD medel dat gebruik is um de euvermaat aon informatie op 'n integraol meneer in de procedures te benutte. 'n Aander winspunt is tot de kalibrasie vaan de zjeometrie vaan de kamera opstèlling tot stand kump door gebruik te maake vaan de metinge zellef zoetot 't gehannes mèt de kalibrasieklómel nimmie nudig is. De veur de kalibrasie nudige stereoskopische pare sjaduwpeziezies weure daobeij otomaties bepaold. Daobeij hoofste vaan de zjeometrie alleijn groof e weg d'n hook tösse de twie kamera-asse te weite.

De in dees dissertasie besjreve methodes zien primair oontwikkelde um gebruik te weure beij 't mete vaan de rezjonaal vervörreming vaan 't hart door 't vollege vaan metaole markeeders in 'n stereoskopische röntgenopstèlling. De methodes zien évels zjinneraoler inzetbaar en weure alleweijl gebruik beij de verwèrreking vaan aofbeeldinge vaan 't minselik hart die verkregen zien mèt d'n MRI tagging techniek en beij 't vollege vaan optische markeeders veur 't mete vaan de rèkdistribusie in skelètspiere.

Curriculum Vitae

11 april 1950	geboren te Maastricht
1962-1968	HBS-B, St. Maartenscollege Maastricht
1968-1972	HTS-Electrotechniek, HTS Heerlen afstudeerrichting Informatietechniek (met lof)
1972-1978	TH-Electrotechniek, TH Eindhoven afstudeerrichting Meet- en Regeltechniek (met lof)
1978-1980	Computer Centrum Limburg, Heerlen Systeemontwerper
1980	Rijksuniversiteit Limburg, Faculteit der Geneeskunde:
1980-1983	Capaciteitsgroep Biofysica. Wetenschappelijk medewerker t.b.v. de statistische ondersteuning van biomedische projecten
1983	Capaciteitsgroep Medische Informatica en Statistiek:
1983-1986	Wetenschappelijk medewerker t.b.v. onderzoek en onderwijs in de Statistiek. Aandachtsgebieden: biomedische toepassingen, longitu- dinale data
1986-1993	Universitair docent t.b.v. onderzoek en onderwijs in de Medische Informatica. Aandachtsgebied: signaal- en beeldverwerking
1993-heden	Vakgroep Medische Informatica. Idem. Aandachtsgebieden: signaal- en beeldverwerking, psychometrische aspecten bij de evaluatie van studieresultaten
1986	Abbott Microcirculation Award
1989	Milieuprijs Gemeente Maastricht

